CONTRIBUTED PAPER

A Multilevel State Estimation Paradigm for Smart Grids

The authors of this paper describe a multilevel state estimator architecture that can sustain growth in size, complexity of data, and information.

By Antonio Gómez-Expósito, *Fellow IEEE*, Ali Abur, *Fellow IEEE*, Antonio de la Villa Jaén, and Catalina Gómez-Quiles, *Student Member IEEE*

ABSTRACT | The main objective of this paper is to describe a multilevel framework that facilitates seamless integration of existing state estimators (SEs) that are designed to function at different levels of modeling hierarchy in order to accomplish very large-scale monitoring of interconnected power systems. This has been a major challenge for decades as power systems grew pretty much independently in different areas, which had to operate in an interconnected and synchronized fashion. The paper initially provides a brief historical perspective which also explains the existing state estimation paradigm. This is followed by a review of the recent technological and regulatory drivers that are responsible for the new developments in the energy management functions. The paper then shows that a common theoretical framework can be used to implement a hierarchical scheme by which even very large-scale power systems can be efficiently and accurately monitored. This is illustrated for substation level, transmission system level as well as for a level between different transmission system operators in a given power system. Finally, the paper describes the use and benefits of phasor measurements when incorporated at these different levels of the proposed infrastructure. Numerical examples are included to illustrate performance of the proposed multilevel schemes.

KEYWORDS | Distributed state estimation; hierarchical state estimation; phasor measurement units (PMUs); smart grids; substation estimation

I. INTRODUCTION

As the penetration of renewable and distributed energy sources along with the necessary means of centralized and distributed energy storage technologies increase to higher and higher levels, the existing electric energy network infrastructures are expected to evolve in two major directions. On the one hand, much longer and higher rated transmission lines carrying both alternating current (ac) and direct current (dc) power at the highest voltage levels will be needed to move huge amounts of renewable energy across very long distances. On the other hand, at the lowest voltage levels, the existing highly centralized power systems will be transformed into the so-called "smart grids" where the intelligence will be to a large extent distributed, through the use of distribution automation, power electronics, active load management, real-time metering, and other technical innovations in telecommunications and computer technologies [1].

State estimators (SEs) determine the most likely state of a power system from sets of remotely captured measurements that are collected periodically by SCADA systems via remote terminal units (RTUs). The role of the SE is crucial in modern energy management systems (EMSs), where a diversity of applications rely on accurate system snapshots [2]. The regulatory wave of the last decade has stressed the importance of the SE tool, in an open-access context in which many more transactions on much more congested networks have to be properly tracked in real time and also recorded for offline engineering studies. This paper serves two purposes. One is to provide a survey of SEs that have been developed and implemented in the past. The other is to present, in a tutorial manner, a novel

Manuscript received August 12, 2010; revised October 19, 2010; accepted January 12, 2011. Date of publication April 29, 2011; date of current version May 17, 2011. The work of A. Gómez-Expósito, A. de la Villa Jaén, and C. Gómez-Quiles was supported by the European Community's 7th Framework Programme (PEGASE Project) under Grant 211407 and the Ministry of Science & Innovation (DGI) under Grants ENE2007-62997 and ENE2010-18867. The work of A. Abur was supported by PSERC/NSF project S-22. The first author received the "VIII Javier Benjumea Puigcerver Research Award, granted by Fundación Focus-Abengoa, for this work.

A. Gómez-Expósito, **A. de la Villa Jaén**, and **C. Gómez-Quiles** are with the Department of Electrical Engineering, University of Seville, Seville 41092, Spain (e-mail: age@us.es; adelavilla@us.es; catalinagq@us.es).

A. Abur is with the Electrical and Computer Engineering Department, Northeastern University, Boston, MA 02115 USA (e-mail: abur@ece.neu.edu).

Digital Object Identifier: 10.1109/JPROC.2011.2107490

multilevel scheme that generalizes a number of previously developed ideas in one comprehensive proposal.

Read from here.

STORICAL PERSPECTIVE

Since its introduction by Schweppe in the late 1960s [3], the SE tool has benefited from a large number of theoretical developments and practical improvements, which were well documented in [4]–[6].

In addition to the pioneering work of Tinney and others on sparsity methods [7], there have been computational improvements such as the so-called fast decoupled state estimator (FDSE), proposed by Monticelli *et al.* [8], based on the successful fast decoupled load flow concept [9]. Another significant improvement was introduced in order to address the issues of numerical stability and ill-conditioning of the conventional weighted least squares (WLS) approach [10]. It was observed that the use of artificially high weights for very accurate measurements such as zero injections and rather low weights for much less accurate pseudomeasurements would lead to poor convergence of the so-called normal equations. It was shown that employing a computationally more expensive method of orthogonal or *QR* factorization would significantly improve the solution accuracy compared to the less stable Cholesky factorization scheme [11]. Alternatively, it was shown that these very accurate or exact measurements, such as zero injections, could be incorporated as explicit equality constraints into the optimization problem formulation, effectively eliminating the need to use artificially high weights [12]. One such formulation used the so-called Hachtel's augmented matrix approach, which could be implemented either directly [13] or using 2×2 pivots in combination with block arithmetic for improved computational efficiency [14], [15].

Network observability and bad data processing constitute two important functions related to the SE problem. Observability analysis is performed in advance in order to determine if the entire state vector is observable and, if not, to identify observable islands. Both numerical [16] and topological [17] algorithms have been proposed and implemented [18]. Strongly related to the network observability analysis was the optimal measurement design that would ensure full network observability under credible loss of RTUs or communication channels [19]. In many instances, the measurement set might be corrupted with gross errors (outliers) and thus, the assumption that all measurement errors were normally distributed would no longer be true. In such cases, if those gross errors (bad data) were not detected and removed by simple plausibility tests before the execution of the SE, the solution would be biased or even the algorithm could fail to converge. Hence, statistical tests such as the chi-squares test and largest normalized residual test, based on chi-squares and standard normal distributions, respectively, were developed in order to detect and identify bad data. Both tests relied on calculated measurement residuals once the SE algorithm converged [20]. More elaborate techniques, such as hypothesis testing identification (HTI), have also been proposed to handle cases involving multiple interacting bad data for which other methods were less effective [21].

There were numerous other developments that addressed a wide spectrum of issues ranging from statistical robustness of estimation [22], [23], hierarchical multiarea estimation [24], inclusion of Ampere measurements at the subtransmission level [25]–[27], incorporation of inequality constraints [28],[29], and detection and identification of network parameter and topology errors [30], [31].

In the last decade there has been an increased interest in the so-called generalized state estimator (GSE), aimed at developing circuit breaker (CB) models to improve the SE capability for topology error processing [5]. This involved a detailed physical level modeling of bus sections that should be used in combination with zooming techniques in order to cope with the huge size of the resulting model [32], [33]. An implicit GSE model has been recently developed. The model maintained the capability to identify topology errors, while using a slightly augmented state vector [34]. The same idea could be used to detect and identify network parameter errors [35].

More recently, the so-called phasor measurement units (PMUs), which provide global positioning system (GPS)-synchronized measurements, among which are voltage and current phasor magnitude and phase angles, are expected to introduce major improvements in SE performance and capabilities [36], [37].

III. EXISTING STATE ESTIMATION PARADIGM

This section briefly reviews the geographical scope, interactions, and solution methodology of existing SE tools.

A. Existing Configuration of SEs

Currently, the scope of SEs is mostly limited to the transmission level, where each transmission system operator (TSO) continuously tracks its own grid from a centralized EMS (see Fig. 1). However, in some power systems, trading and pricing may also take place at lower voltage levels, which are typically not closely monitored by a SE. This is mainly due to the unavailability of sufficiently redundant set of measurements at these voltage levels. Hence, there is a need to improve the monitoring capability for these parts of the systems and to facilitate reliable and effective operation of power markets at these levels as well.

At the uppermost voltage levels, where interconnections for energy trading exist, tie-line power flows have to be properly monitored, for which TSOs must get access to both the electrical model and real-time measurements of its neighbors, at least those in the adjacent buses. For this

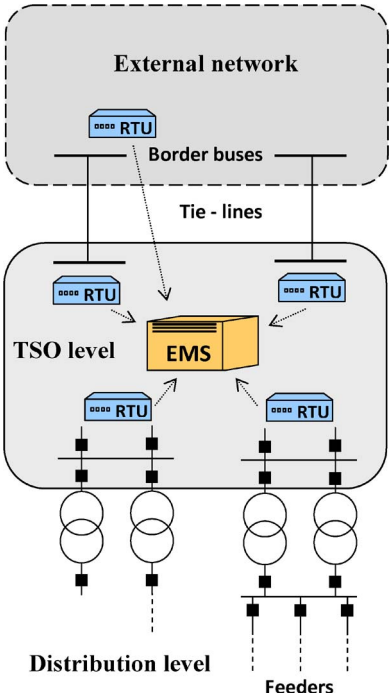


Fig. 1. *Conventional TSO-level state estimation paradigm. (RTU: remote terminal unit; EMS: energy management system; TSO: transmission system operator.)*

purpose, the external grid is usually represented by a reduced equivalent circuit [38]. Therefore, existing SEs, being tailored to the needs of a single TSO, with very neat borders, are not designed to significantly interact with its neighbors or subordinate networks. As explained later, this is not a satisfactory state of affairs, in view of both the needs and possibilities offered by smart grids, but commercial software evolves at a lower pace than theoretical developments.

B. Standard WLS-Based Solution

The SE relies on the following measurement equation [3]:

$$z = h(x) + e \tag{1}$$

where x is the state vector to be estimated (size n), z is the known measurement vector (size $m > n$), h is the vector of functions, usually nonlinear, relating error free measurements to the state variables, and e is the vector of measurement errors, customarily assumed to have a normal distribution with zero mean and known covariance matrix R. When errors are independent, R is a diagonal matrix with values σ_i^2, where σ_i is the standard deviation of the error associated with measurement i.

In conventional bus-branch SE models the state vector is composed of voltage magnitudes and phase angles, whereas the measurement vector typically comprises power injections, branch power flows, and voltage magnitudes. Recently, the availability of synchro-phasors (PMUs) has made it possible the incorporation of phase angle measurements into the SE process.

The WLS estimator minimizes the weighted squares of residuals of the measurements given by

$$J = \sum_{i=1}^{m} W_i r_i^2$$

where $r_i = z_i - h_i(\hat{x})$ is the measurement residual, $\hat{x}$ is the estimated state vector, and W_i is the respective weighting coefficient.

The state estimate can be obtained by iteratively solving the normal equations

$$G_k \Delta x_k = H_k^T W [z - h(x_k)] \tag{2}$$

where $H_k = \partial h / \partial x$ is the Jacobian evaluated at $x = x_k$, $G_k = H_k^T W H_k$ is the gain matrix, $W = R^{-1} = \mathrm{diag}(W_i)$ is the weighting matrix, and $\Delta x_k = x_{k+1} - x_k$, k being the iteration counter.

Iterations are terminated when an appropriate tolerance is reached on Δx_k. The covariance of the estimate is

$$\mathrm{cov}(\hat{x}) = G_k^{-1}.$$

Then, the bad data processing function, aimed at detecting, identifying, and eliminating bad analog measurements, is activated. This is accomplished through the largest normalized residual test [20].

IV. TECHNOLOGICAL AND REGULATORY DRIVERS

The SE paradigm described above will have to drastically change with the advent of the smart grid. On the one hand, new generations of digital devices, such as PMUs or intelligent electronic devices (IEDs), intended for measurement, protection, and control, being less expensive and more flexible than existing analog equipment, will invade virtually every corner of future networks [37], [39]. This will provide a more accurate, complex, and highly redundant information system, allowing SEs to extend their scope well beyond presently observable areas and also to incorporate advanced functions that have not yet reached the industrial stage, in spite of being conceptually mature on the researcher blackboard. On the other hand, smart

transmission grids should further promote the development of regional energy markets, involving distant energy transactions. This also implies wide-area physical interactions, possibly of catastrophic consequences in case of cascaded failures [40], [41].

In the following sections, outstanding technological innovations associated with the smart grid concept will be succinctly analyzed from the point of view of their influence in the conceptual design of future SE architectures.

A. Intelligent Electronic Devices

The protection, metering, and control functions in substations are naturally distributed by the role and location of each device, being designed in general to provide primary protection or monitoring of an individual substation equipment. These functions may be performed by smart multifunctional and communicative units, so-called IEDs. They are broadly defined in [42] as "devices incorporating one or more processors with the capability to receive or send data/control signals from or to an external source (e.g., electronic multifunction meters, digital relays, controllers)."

The IEDs, employing efficient signal processing techniques, are becoming the source of much more information in real time than the one existing in old substations. Apart from implementing specific protection or control algorithms, they can provide externally electrical magnitudes measured by protection transformers as well as phase differences among them [39]. Those measurements can be synchronized, both at the substation and wide-area levels, by means of the GPS satellite clock time reference.

The quality of the SE process strongly relies on the redundancy of the measurement set. For this reason, the possibility of incorporating all the information provided by IEDs, including the ones whose primary function is not measurement but protection, is very attractive.

B. Communication and Architecture Standards: IEC 61850

New system architectures need to be devised for preprocessing the ever-increasing amount of information gathered by the IEDs. This goal is achieved by replacing the conventional centralized systems, based on RTUs and numerous protection and control devices, with local-area-network-based systems and advanced multifunctional protection and control IEDs [43].

IEC 61850, the global communication standard for substation automation system, defines the communication between IEDs and not only solves the interoperation problem but also specifies other system requirements, like message performance and information security in substation automation system network [43]. IEC 61850 allows

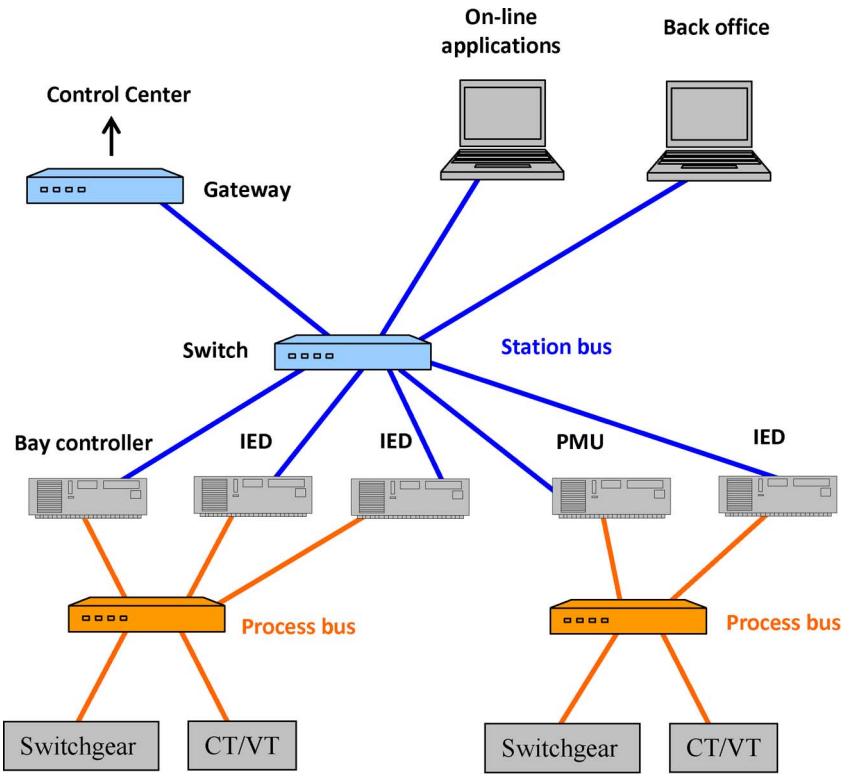


Fig. 2. *Substation-level hardware platform, according to IEC 61850. (IED: intelligent electronic device; PMU: phasor measurement unit; CT/VT: voltage/current transformer.)*

interoperability of IEDs from different manufacturers without the use of protocol converters.

The standard defines two communication buses between the different subsystems within a substation (see Fig. 2). The process bus is devoted to gathering information about electrical magnitudes, such as voltage or current, as well as switching status information, from the transformers and transducers connected to the primary power system process. The station bus is aimed at allowing primary communications between the logical nodes, which provide the various station protection, control, monitoring, and logging functions. The communication technologies involved in these buses include: Ethernet on fiber optic, TCP/IP, and MMS (ISO9506). This architecture supports remote network access for all types of data reads and writes.

C. Phasor Measurement Units

Most of the EMS applications assume that the system is in a pseudosteady state where ac circuit analysis can be carried out using phasors. Network model and the voltage phasors at all system buses are used to determine the state of the system. While the bus voltage phasors can be estimated based on a redundant set of measurements, direct measurement of phasors will be possible only if measurements are synchronized. PMUs are devices that take advantage of the GPS satellites in order to time synchronize the measurements. Voltage and current signals are collected at secondaries of instrument transformers (CT and PT) and are sampled via analog-to-digital converters at 48 samples/cycle. These samples are then processed and synchronized with universal time coordinated (UTC) time from a GPS receiver within 1-μs accuracy [36]. Time-synchronized samples are processed to obtain time-stamped voltage and current phasors, which are then transmitted over Ethernet to phasor data concentrators (PDCs), which will then send data to control center SCADA server. IEEE Standard C37.118 describes the requirements, format, and communication protocol for data provided by PMUs [44].

PMUs may have several channels, each of which will record one phase of a voltage or current signal. Two sets of three channels are typically used for three phase voltage measurements at the substation and several sets of three channels will be used for measurement of three phase currents along incident lines or transformers. Positive sequence components rather than individual phase signals are typically used by network applications, hence three phase signals are processed to compute the positive sequence components. The string of positive sequence phasors that are computed by the PMU will then be communicated at 30 samples/s to the PDC.

D. Distribution Automation

Distribution automation is a mature concept whose real potential never took off by a lack of reasonably priced infrastructures. In fact, unlike in transmission systems, most functions in this field (fault detection, service restoration, network reconfiguration, etc.) have been traditionally performed with the help of mobile service teams on call.

Regarding the information that can be found at the distribution level, virtually no measuring devices have been installed until recently to monitor the operating condition of medium voltage (MV) feeders. Typically, the current (or sometimes the power flow) at each feeder head, along with the voltage magnitude at the MV bus, are telemetered and gathered at dedicated distribution management systems (DMSs). But no real-time information is obtained of what happens downstream, unless a fault occurs. The substation bus voltage is kept almost constant by the use of automatic under-load tap changers, in the hope that customer voltages remain acceptable for nearly all operating points.

This situation is rapidly changing for several reasons.

- Distributed generators (DGs) connected at this level frequently reverse the sign of power flows, creating overvoltage problems that should be properly monitored and prevented. On the other hand, the energy they inject, frequently at premium prices, should also be carefully monitored and recorded.
- Smart meters, currently being deployed, provide hourly customer demands via PLC, regular cellular phone technology, or alternative means. There is a trend to concentrate all this information at the secondary transformer centers, from where it is then submitted upstream to the distribution substation or DMS.
- Cheap fault current detectors, along with automatic or remotely operated reclosers, are being installed at strategically selected points to speed up the service restoration process. These devices are capable of providing less accurate current values that could also be attached to the remaining information sent to the substation.

Consequently, the massive introduction of DG and a plethora of distribution automation devices, at network levels that are not currently supervised by TSOs, will also contribute to the development of ubiquitous monitoring systems [45].

E. Wide-Area Regional Energy Markets

Energy markets outcomes may be significantly affected by unique information regarding the present and likely states of the grid. But gathering such information is a real challenge when the energy transactions take place over networks that cross national or regional market borders. For this and other strategic reasons, international regulatory entities are promoting worldwide the creation of regional-level system operators, in an attempt to eliminate barriers and better coordinate multi-TSO transactions.

In the United States, for instance, realizing that competition was hindered because only a handful of utilities owned and controlled a large portion of the region's transmission, the Federal Energy Regulatory Commission (FERC) issued the Order 890 "Preventing Undue Discrimination and Preference in Transmission Service" and the 2010 Notice of Proposed Rulemaking on "Transmission Planning and Cost Allocation by Transmission Owning and Operating Public Utilities." These orders amended its regulations in order to remedy opportunities for this undue discrimination and address deficiencies in the proforma open-access transmission tariff. Enabling interconnection-wide operation via these rules that extend the local, regional, and inter-regional planning processes will be greatly facilitated by the proposed multilevel state estimation scheme.

Across the Atlantic, the European Regulators' Group for electricity and gas (ERGEG) launched in 2007 an initiative to create a series of Regional Energy Market projects within the European Union (EU), in order to remove barriers to cross-border trade between countries as a first step towards the completion of a single EU market for electricity [46]. Other long-term ambitious projects for long-distance bulk transmission of renewable energy, such as the Desertec initiative [47], will further stress the need for transnational cooperation in network monitoring.

V. MULTILEVEL STATE ESTIMATION ARCHITECTURE

The technological developments discussed above, along with more aggressive regulatory schemes intended to promote efficient trading of clean energy, allow to envision a future in which SEs will spread from MV distribution feeders to the bulk extra high voltage (EHV) transmission network, spanning several interconnected areas [45], [48].

But the crucial question is how the existing SE paradigm will have to be adapted in order to cope with such a diverse and extensive geographical scope as well as the formidable amount of information provided by the heterogeneous and distributed sources arising in the upcoming smart grid environment. When trying to answer that question, a dilemma arises about whether it will be feasible and convenient to keep on submitting all this information to a central EMS or it should be processed to a large extend in a local manner, as close as possible to the place in which it is generated. The first choice is discouraging for two main reasons:

1) the investment in new communication infrastructures would be prohibitive;
2) the required computing power for the central entity to be capable of processing the incoming data in real time should be one or two orders of magnitude larger than that of existing EMS.

The alternative and most natural choice to deal with the explosion of information arising in this multiagent

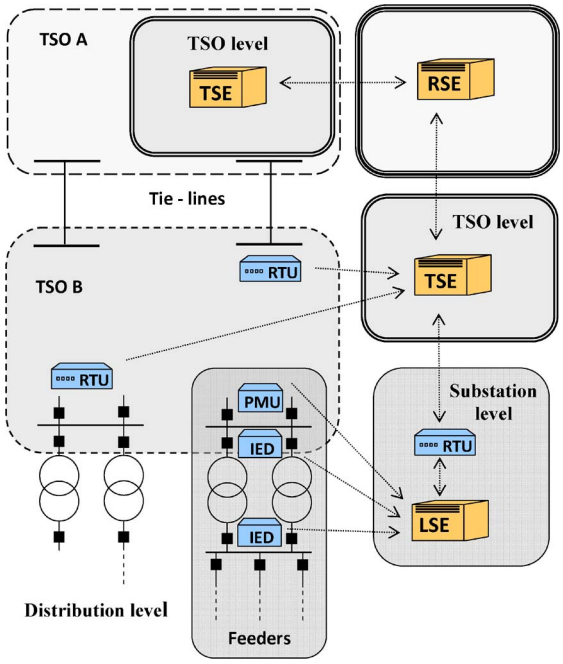


Fig. 3. *Smart-grid-oriented multilevel state estimation paradigm. (LSE: local SE; TSE: transmission-level SE; RSE: regional multi-TSO SE.)*

distributed environment calls necessarily for a multilevel, hierarchical SE paradigm. In this paper, at least three major levels are identified, as shown schematically in Fig. 3 (a fourth level associated with distribution feeders will also be distinguished).

At the lowest level, a local SE (LSE) can be implemented to preliminarily deal with the information collected within a substation or small set of adjacent substations. A great majority of raw measurements will be processed at this distributed level, where a modest but sufficient computing power already exists. Distribution substations, delivering power to a large number of secondary transformers through a set of radial feeders, constitute a particular relevant case. In those substations, it is advantageous and makes sense to process each radial feeder in a decoupled manner, leading to a fourth level of information processing.

The results provided by the LSE have to be transmitted through existing RTUs and communication channels to the TSO-level SE (TSE). At this intermediate level, commercially available software can be adopted with minor modifications, the major difference with respect to a conventional SE being that prefiltered rather than raw measurements are handled.

At the uppermost level, a regional SE (RSE) will be needed to synchronize and refine the results separately provided by each TSO affiliated with the interconnected system, particularly near the border nodes. The RSE will be a customized tool, designed in such a way that the amount of information exchanged with subordinate TSEs

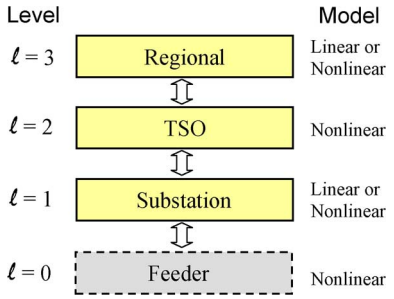


Fig. 4. *Hierarchical multilevel architecture.*

is kept to a minimum. This SE level will significantly benefit from wide-area measurements provided by PMUs.

Fig. 4 shows the resulting hierarchy, including the feeder level arising in distribution substations. The double arrow represents schematically the bidirectional flow of information between adjacent levels.

The interactions between adjacent levels can be better formulated and justified as particular customized cases of a common theoretical framework that will be presented in the next section. Then, a more detailed treatment of each level will be separately made in the remaining sections.

VI. THEORETICAL FRAMEWORK

The multilevel SE formulation advocated in this paper requires that the standard WLS solution approach be reconsidered. This stems from the following observations.

- The conventional TSE is based on the so-called bus-branch electrical model, which is not of direct application to the substation level. In this environment, the LSE must be capable of dealing with extra raw measurements and detailed topology information, not found ordinarily at the EMS level, for which augmented state vectors must be considered.
- The estimate of the multilevel SE should be optimal, i.e., as close as possible to that provided by an ideal solver simultaneously handling all raw measurements for the entire set of interconnected areas. Theoretically, this is possible, but only if the required statistical information, associated with each partial estimate, is duly exchanged between adjacent levels, usually in an iterative manner. Existing TSEs should be adapted to interact in this way with their neighbors.

The conventional WLS-based SE methodology, currently used by TSOs worldwide, was succinctly reviewed in Section III-B. In what follows, a recently introduced, two-stage WLS solution method, based on a factorization of the measurement model, will be presented [49]. Such a scheme, when combined with suitable geographical decomposition techniques and certain model transformations, determines the algorithmic steps and interactions involved in the multilevel SE paradigm, and provides the right mathematical framework supporting the overall estimate optimality.

First, the two-step approach introduced in [49] will be generalized to the fully nonlinear case. Then, two relevant particular cases, in which one of the resulting submodels is linear, will be considered. The final subsection provides important implementation guidelines allowing the factorized solution to be geographically distributed whenever possible, which is the main goal of the proposed multilevel architecture.

A. Factorized WLS Solution: General Nonlinear Case

The factorized approach to solve the WLS problem arises when the nonlinear measurement model (1) is "unfolded" into two sequential WLS problems, as follows:

$$z = f_1(y) + e \tag{3}$$
$$y = f_2(x) + e_y \tag{4}$$

where y is a vector of intermediate variables, selected in such a way that the solution of the pair (3)–(4) offers an advantage over that of (1). For the resulting factorized model to be equivalent to the original one, the following condition must be satisfied:

$$h(x) = f_1[f_2(x)] \quad \Rightarrow \quad H(x) = F_1(y)F_2(x) \tag{5}$$

where H, F_1, and F_2 represent the Jacobian matrices of h, f_1, and f_2, respectively.

As shown in the Appendix, the optimal estimate $\hat{x}$ provided by the conventional iterative process (2) can be alternatively obtained by successively solving the following pair of equations:

$$\left[F_1^T W F_1\right]\Delta y_k = F_1^T W[z - f_1(y_k)] \tag{6}$$
$$\left[F_2^T G_1 F_2\right]\Delta x_k = F_2^T G_1[\tilde{y} - f_2(x_k)] \tag{7}$$

where $\tilde{y}$ in (7) is the estimate of the intermediate vector provided by (6), and the weighting matrix G_1 satisfies

$$G_1 = [\mathrm{cov}(y)]^{-1} = F_1^T W F_1. \tag{8}$$

In other words, the weighting matrix of the second WLS problem is the gain matrix of the first one.

As explained in the Appendix, full equivalence between the original and the factorized models requires that the linearization of f_1 and f_2 be performed at the same point in

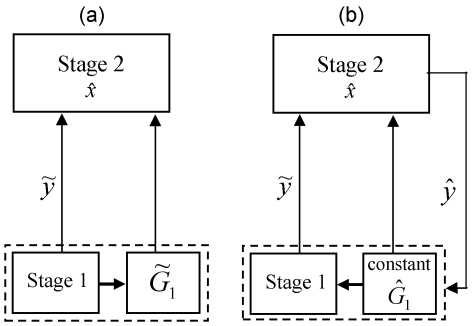


Fig. 5. *Schematic flowchart of the two-stage factorized procedure.*

the n-dimensional space represented by x. This leads in the general case to an outer iterative process in which two stages, one for each WLS subproblem, are successively solved.

Accordingly, the factorized procedure can be formally decomposed into the following sequence of steps, where the first run is separately considered for clarity of presentation.

First run:

- *Stage 1*: Find $\tilde{y}$ by repeatedly solving (6) until convergence. As a byproduct, $\tilde{G}_1$ is available.
- *Stage 2*: Find $\hat{x}$ by repeatedly solving (7) with $G_1 = \tilde{G}_1$.

Subsequent runs (if needed):

- *Stage 1*: Update the Jacobian, $\hat{F}_1$, and the gain matrix, $\hat{G}_1$, for $\hat{y} = f_2(\hat{x})$. Keeping these matrices constant, find $\tilde{y}$ by repeatedly solving (6).
- *Stage 2*: Find $\hat{x}$ by repeatedly solving (7) with $G_1 = \hat{G}_1$.

The process is repeated until two consecutive runs of Stage 2 provide close enough values of $\hat{x}$.

Fig. 5 represents schematically the two-stage procedure outlined above. The auxiliary vector y plays the role of a state vector when solving Stage 1 and that of a measurement vector when solving Stage 2. Also, as stated above, the gain matrix of Stage 1 becomes the weighting matrix of Stage 2. Note that this matrix is no longer diagonal, but its sparsity can and should be fully exploited. The fact that subsequent runs of Stage 1 involve only constant matrices can also be exploited.

In practice, unless the measurement vector is very noisy and/or contains key bad data, the first run of stages 1 and 2 will provide sufficiently accurate results.

For the sake of clarity, the description provided in this section assumes that all raw measurements z can be used in Stage 1. If this is not the case, then Stage 2 can be easily redesigned to handle simultaneously the pseudomeasurement $\tilde{y}$ and the components of z not yet used during Stage 1. This is explained in detail in [49].

The fully nonlinear factorized model presented above may reduce to any of the two particular cases discussed below.

B. Factorized WLS Solution: Linear–Nonlinear Case

This case, by far the most interesting in practice, arises when the intermediate vectory can be chosen in such a way that the measurement model of Stage 1 becomes linear [49]. Then, the nonlinear systems (3) and (4) reduce to

$$z = Ay + e \tag{9}$$

$$y = f_2(x) + e_y. \tag{10}$$

The general nonlinear model can be easily particularized to this case by systematically replacing $f_1(y)$ and F_1 in the expressions above by Ay and A, respectively. The most relevant implication is that the iterative system (6) reduces in this case to the following linear one:

$$[A^T W A]\tilde{y} = A^T W z. \tag{11}$$

Accordingly, the gain matrix

$$G_1 = A^T W A \tag{12}$$

remains constant, so long as the network topology and measurement set structure are unaltered. Therefore, changes in the measurement values and state variables originated by the daily load evolution do not alter the linear model of Stage 1, which is one of the sources of computational saving associated with the factorized approach.

The factorized procedure reduces in this case just to the first run of Stages 1 and 2. Stage 1 constitutes a linear prefilter of the raw measurement vector, which can be a valuable tool by itself (for instance, to perform preliminary bad data analysis when the redundancy if sufficiently high). Accuracy of the linear estimate delivered by Stage 1 could be subsequently checked by comparing $\tilde{y}$ with $\hat{y} = f_2(\hat{x})$.

C. Factorized WLS Solution: Nonlinear–Linear Case

In this case, the model of Stage 2 becomes linear. This may arise when all raw measurements are processed during Stage 1 and no lossy network components are involved in Stage 2, which reduces typically to a trivial model relating state variables with their estimates provided by Stage 1 [50]. Then, the nonlinear systems (3) and (4) reduce to

$$z = f_1(y) + e \tag{13}$$

$$y = Bx + e_y. \tag{14}$$

The general model can be easily particularized to this case by systematically replacing $f_2(x)$ and F_2 by Bx and B,

respectively. The most relevant implication is that the iterative system (7) reduces to the following linear one:

$$[B^T G_1 B]\hat{x} = B^T G_1 \tilde{y}. \tag{15}$$

Note that Stage 2 involves a constant, usually quite sparse, and trivial Jacobian, B, but a weighting matrix G_1, which compactly embeds most of the information associated with the original SE problem (raw measurement covariance, network topology, and parameters).

Like in the general nonlinear case, the factorized scheme may theoretically involve in this particular case several executions of Stages 1 and 2. In practice, however, it is seldom needed to repeat the two-stage process, unless the required accuracy is extremely high or the raw measurement set is abnormally noisy. A computationally less expensive alternative when subsequent solutions are needed consists in updating only the right-hand side of (15). This way, the *LU* or *QR* factors of the coefficient matrix do not have to be recomputed.

D. Distributed Implementation of the Two-Stage Procedure

The two-stage SE procedure, in the basic form outlined above, can be a better choice than the conventional scheme (particularly in the linear–nonlinear case) provided a suitable set of intermediate variables y can be found. The improved performance may take the form of reduced computational effort (simpler and/or constant Jacobian components, smaller size of the state vector during the iterative process), early bad data processing capability, simpler observability analysis, etc. [51].

However, for the multilevel paradigm envisioned in this paper, it is most important to consider a distributed environment, in which geographically scattered measurements can be naturally grouped in clusters, weakly coupled with neighboring sets. Fig. 6(a) illustrates this common situation for the case of four areas or clusters. Within each area, two types of variables can be distinguished: 1) internal variables, not involved in the measurement models of neighbor areas; 2) border variables, appearing in the measurement model of at least an adjacent area. Accordingly, the measurement model (1) can be decomposed in the following manner for the case of j clusters:

$$\begin{aligned} z_1 &= h_1(x_{i1}, x_b) + e_1 \\ z_2 &= h_2(x_{i2}, x_b) + e_2 \\ &\vdots \\ z_j &= h_j(x_{ij}, x_b) + e_j \end{aligned} \tag{16}$$

where x_{ik} represents the set of internal variables for area k and x_b comprises the union of all border variables. Ideally,

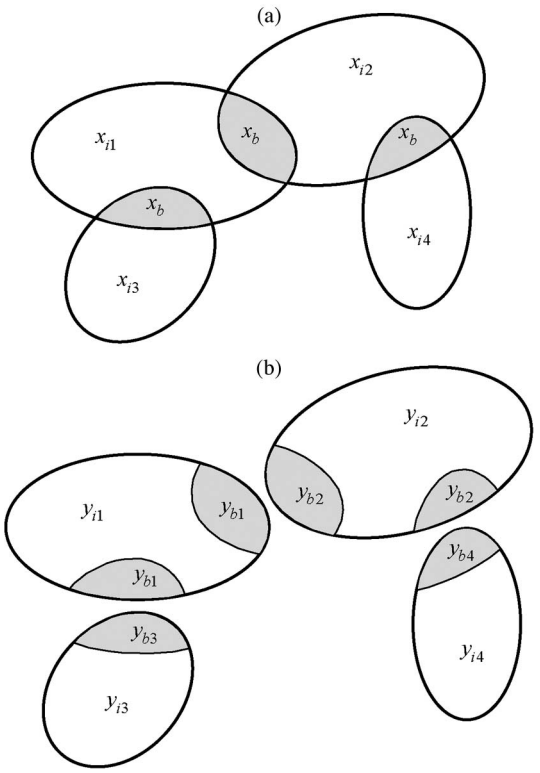


Fig. 6. *Measurement model composed of four natural clusters.*

the number of interior variables should be much larger than that of border variables, but in practice this may not be always the case. Notice that the set x_b constitutes the coupling term among all measurement submodels.

In this context, as justified in Section V, it is advantageous (sometimes even mandatory) to process the raw information as close as possible to the level (feeder, substation, etc.) in which it is captured, usually in a decoupled fashion. This calls for a distributed implementation of Stage 1, the aim of Stage 2 being essentially to coordinate or refine the preliminary solution provided by Stage 1 [49].

Keeping this goal in mind, the intermediate vectory should be selected in such a way that Stage 1 reduces to a set of fully decoupled SE subproblems, as shown in Fig. 6(b). Unlike x_b, whose components are shared by two or more areas, the vector of border variables y_b is composed of disjoint components, each involved in the measurement model of a single cluster. Depending on the measurement model and spatial structure of the particular SE problem being decomposed, the augmented vector y_b can be obtained by simply replicating border state variables of the original problem, adding certain measured magnitudes to the state vector, etc.

According to the geographical decomposition achieved, the measurement vector z is split into j components, each one exclusively related with the respective components of y. Therefore, the decoupled measurement

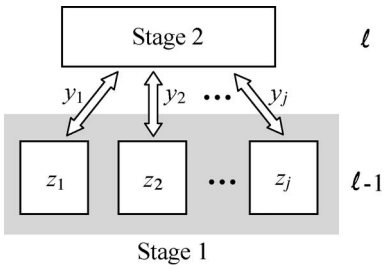


Fig. 7. ***Distributed implementation of Stage 1.***

models of Stage 1 can be mathematically formulated as follows:

$$\begin{aligned} z_1 &= f_{11}(y_{i1}, y_{b1}) + e_1 \\ z_2 &= f_{12}(y_{i2}, y_{b2}) + e_2 \\ &\vdots \\ z_j &= f_{1j}(y_{ij}, y_{bj}) + e_j. \end{aligned} \tag{17}$$

Fig. 7 schematically illustrates the interactions between both stages in case Stage 1 is performed in a distributed manner. The solution corresponding to each cluster can be iteratively obtained, in the nonlinear case, by solving the associated normal equation system (6).

Regarding Stage 2, consider for simplicity the nonlinear–linear case (extending the distributed formulation to the fully nonlinear context is straightforward). Let $\tilde{y}_b$ and $\tilde{y}_i$ be the estimates provided by Stage 1 of the border and interior variables, respectively, for all clusters $1, 2, \ldots, j$. Then, the mathematical model of Stage 2, for the linear case, can be written as follows:

$$\tilde{y}_b = Bx_b + e_b \tag{18}$$
$$\tilde{y}_i = x_i + e_i. \tag{19}$$

The weighting matrix arising in Stage 2 is a block diagonal matrix formed by simply juxtaposing the individual gain matrices corresponding to each cluster of Stage 1, yielding

$$G_1 = \begin{bmatrix} G_{11} & & & \\ & G_{12} & & \\ & & \ddots & \\ & & & G_{1j} \end{bmatrix} = \begin{bmatrix} \blacktriangle & \diamond & & & & & & \\ \diamond & \blacktriangledown & & & & & & \\ & & \blacktriangle & \diamond & & & & \\ & & \diamond & \blacktriangledown & & & & \\ & & & & \ddots & & & \\ & & & & & & \blacktriangle & \diamond \\ & & & & & & \diamond & \blacktriangledown \end{bmatrix} \tag{20}$$

where, as suggested by the rightmost matrix structure, each diagonal block is composed of four submatrices (self and mutual covariance terms between interior an border variables).

Symmetrically reordering the rows/columns so that border and interior variables are grouped together leads to the following blocked structure:

$$G_1 = \begin{bmatrix} G_{bb} & G_{ib} \\ G_{ib}^T & G_{ii} \end{bmatrix} = \left[\begin{array}{cc|cc} \blacktriangle & & \diamond & \\ & \blacktriangle & & \diamond \\ & \diamond & \blacktriangledown & \\ \diamond & & & \blacktriangledown \end{array}\right] \tag{21}$$

Each and every major block above is composed of j decoupled blocks. Furthermore, the block sizes in G_{ii} (bottom right) are frequently much larger than those of G_{bb} (top left), which is an important feature from the computational point of view (e.g., partly distributed solution of Stage 2).

Based on the above notation, the normal equations to be solved at Stage 2 can be formulated

$$\begin{bmatrix} B^T & \\ & I \end{bmatrix} \begin{bmatrix} G_{bb} & G_{ib} \\ G_{ib}^T & G_{ii} \end{bmatrix} \begin{bmatrix} B & \\ & I \end{bmatrix} \begin{bmatrix} x_b \\ x_i \end{bmatrix} = \begin{bmatrix} B^T & \\ & I \end{bmatrix} \begin{bmatrix} G_{bb} & G_{ib} \\ G_{ib}^T & G_{ii} \end{bmatrix} \begin{bmatrix} \tilde{y}_b \\ \tilde{y}_i \end{bmatrix} \tag{22}$$

and, rearranging

$$\begin{bmatrix} B^T G_{bb} B & B^T G_{ib} \\ G_{ib}^T B & G_{ii} \end{bmatrix} \begin{bmatrix} x_b \\ x_i \end{bmatrix} = \begin{bmatrix} B^T G_{bb} & B^T G_{ib} \\ G_{ib}^T & G_{ii} \end{bmatrix} \begin{bmatrix} \tilde{y}_b \\ \tilde{y}_i \end{bmatrix}. \tag{23}$$

The above linear system could be directly solved by any of the generic techniques developed for symmetric systems (e.g., Cholesky or orthogonal factorization). However, this would be wasteful for two main reasons: 1) no advantage is taken of the distributed nature of the problem, reflected by the blocked structure of the coefficient matrix; 2) no benefits are obtained from the existing factorization of G_1, performed during Stage 1.

Many published procedures on multiarea SE ignore the mutual covariance between internal and border variables ($G_{ib} \approx 0$), as a means of quickly obtaining a suboptimal

solution. In fact, crude diagonal approximations of the covariance matrix components can also be found when formulating the coordination problem [52], [53]. As will be seen in the tutorial examples presented later, arbitrarily neglecting mutual covariance values may lead to poor solutions. Therefore, future SE implementations should be redesigned to account for nondiagonal weighting matrices, which is the price paid for the strengthened interactions of existing SEs with their neighborhood.

There exist at least two alternatives to more efficiently solve the system (23), taking into account the specific structure of the problem, that will be briefly discussed in the following.

1) Elimination of x_i: After trivial algebra, the system (23) can be reduced to

$$(B^T G_{\text{sch}} B) x_b = B^T G_{\text{sch}}\, \tilde{y}_b \tag{24}$$

where the resulting coefficient matrix

$$G_{\text{sch}} = G_{bb} - G_{ib} G_{ii}^{-1} G_{ib}^T$$

is known as the Schur's reduction of G_1.

The following remarks are in order regarding this solution approach.

- The vector component $\tilde{y}_i$, comprising the sets of interior variables for all clusters, is missing in (24), which is a nice feature. Note, however, that the influence in Stage 2 of the interior components of the solution provided by Stage 1 is properly exerted through the mutual covariance matrix G_{ib}.
- When computing G_{sch}, advantage should be taken of the block diagonal structure of G_1 components, as shown by (21).
- The inverse of G_{ii} is never computed, but the available sparse factors of its j block components should be instead resorted to. Therefore, computing G_{sch} reduces basically to repeated solutions of the sparse linear systems arising in Stage 1, which can be fully distributed.

2) Block Gauss–Seidel Iterations: The system (23) can be also rewritten as two coupled subsystems

$$(B^T G_{bb} B) x_b = B^T G_{bb}\, \tilde{y}_b - B^T G_{ib} (x_i - \tilde{y}_i) \tag{25}$$

$$G_{ii} (x_i - \tilde{y}_i) = G_{ib}^T (\tilde{y}_b - B x_b). \tag{26}$$

Then, a block iterative scheme can be implemented as follows.

1) Initialization: $x_i \leftarrow \tilde{y}_i$.
2) Obtain x_b by solving (25) with the most recent value of x_i.
3) Using x_b from the previous step, update x_i by solving (26).
4) Repeat steps 2) and 3) until convergence.

The computational effort and the amount of information exchanged is significantly reduced if step 3) above is fully distributed among the j processors in charge of Stage 1. In this regard, it is more convenient for the auxiliary vector

$$c_{ib} = G_{ib} (x_i - \tilde{y}_i) \tag{27}$$

to be locally computed and sent to the central processor, instead of x_i, in order to perform step 2).

Note that, when the iterative process is initialized in this way, the first value of x_b is the same as that obtained by neglecting G_{ib} in (23).

The former alternative, based on the reduced form, will be preferable in general when the number of border variables is comparatively small and/or the number of iterations for the latter scheme is high.

In the following sections, the two-stage model, information exchanged, and main features of the SE levels previously identified will be presented. A small example will be included illustrating each case. The basic idea is to consecutively apply the distributed version of the two-stage procedure (Fig. 7) to each pair of adjacent SE levels (Fig. 3).

VII. LOCAL STATE ESTIMATION

As discussed above, the advent of IED and PMU technologies at the substation level, along with the associated communication and computing infrastructures, is paving the way to the future "smart substation" [54]. Fig. 2 shows in schematic form the main actors arising in this distributed architecture, according to the IEC 61850 standard. This covers all aspects of the communications between a network of devices in the substation and the related systems.

In this environment, a huge number of measurement points, including those associated with protective devices, will provide information at a high sampling rate, which calls for the implementation of a LSE.

The LSE is useful not only to prefilter and reduce the size of the measurement set that should be sent to the TSE, but in many instances also for early detection of model and network inconsistencies at the substation level. For the redundancy ratios expected in next-generation substations, most topology errors and bad data could be handled at this level [55].

Several cases can be considered, depending on the substation voltage level and type.

A. Transmission Substation: Linear Model

This case arises by grouping together, in a single cluster or area, all busbar sections corresponding to the same rated voltage in a substation, along with the associated switching, protective and measurement devices [49]. A single substation then gives rise to as many areas as voltage levels. Fig. 8 shows a real substation with two voltage levels, each one leading to an area that is separately processed by the LSE.

An area so defined is therefore connected by nonzero impedance lossy branches (lines and power transformers, referred to hereafter as external branches) to its neighbors. On the other hand, internal connections take place exclusively through lossless components, leading to a straightforward linear model. This model, composed in turn of three decoupled submodels (active and reactive power and voltage magnitude), can be systematically built by resorting to well-known topological properties and concepts [56]. In the presence of PMU measurements, another submodel for phase angles could be considered.

Based on the so-called "proper tree," and the resulting links, a set of state variables can be chosen [34]. For convenience, the tree is selected in such a way that power flows through external branches are contained in the state vector. Those power flows, along with the voltage magnitudes of candidate electrical buses, constitute the border variables of each area or cluster (y_b). The power flows through the remaining links, necessary to complete the state vector, define the interior component (y_i).

A detailed model, relating existing measurements and topological constraints with the local state vector y, can be established for each area. Then, if sufficient redundancy locally exists, an estimate $\tilde{y}$ is provided by solving (11). The associate covariance matrix can also be obtained.

To illustrate the above ideas, consider for instance area 1 in Fig. 8, in which all CBs are closed. The associate

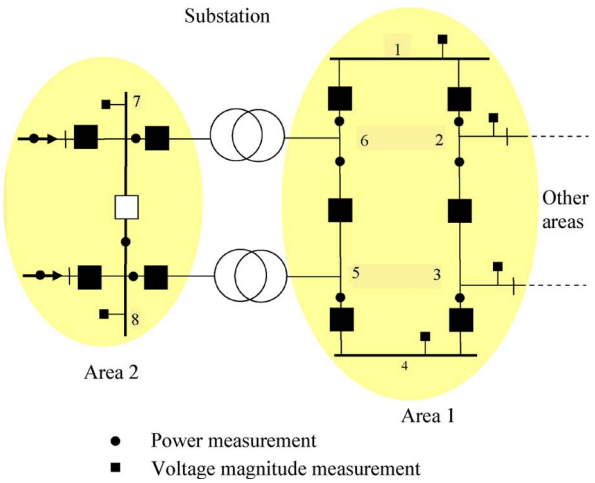


Fig. 8. *Sample substation and associate measurements.*

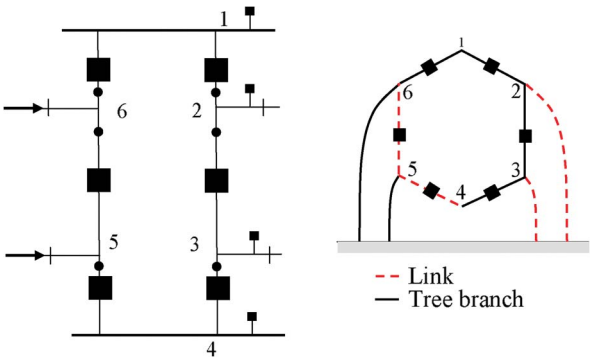


Fig. 9. *Sample linear substation and associate graph.*

graph and a suitable tree are shown in Fig. 9 (for simplicity, it is assumed that only buses 2 and 3 are connected to the rest of the TSO network, while buses 5 and 6 are representing external injections).

From the perspective of the two-stage procedure, in which the substation is considered a component (Stage 1) of a larger TSO system (Stage 2), the state vector can be decomposed as follows:

$$y_b^T = [P_{2e}, P_{3e}, Q_{2e}, Q_{3e}, V_1]$$
$$y_i^T = [P_{45}, P_{56}, Q_{45}, Q_{56}].$$

To estimate the above set of nine variables, 16 measurements are available, 12 power flows, and four voltage magnitudes (the measurement points correspond with those of a real substation).

The reader can easily verify that any measurement is a linear combination of the state variables. For instance, the expression for measurement P_{12} is

$$P_{12}^m = P_{2e} + P_{3e} + P_{45} + e_{p12}$$

and that of V_4

$$V_4^m = V_1 + e_{v4}.$$

Systematically gathering all the required relationships leads to a linear system of the form (9), which is composed in this case of three decoupled subsystems (P, Q, and V).

It is worth noting that the substation graph comprises just four loops, which means that the null-injection constraints P_1 and P_4 are implicitly considered when defining the graph [34], [56].

Table 1 Values of State Variables for the Linear Substation Case

		Exact	Estimated
y_b	P_{2e}	1.3000	1.3300
	P_{3e}	1.1000	1.0700
	Q_{2e}	0.1300	0.1265
	Q_{3e}	0.1100	0.1130
	V_1	1.0200	1.0180
y_i	P_{45}	-1.2000	-1.1900
	P_{56}	0.2000	0.1800
	Q_{45}	-0.1200	-0.1200
	Q_{56}	0.0200	0.0180

For a given measurement snapshot, generated by adding Gaussian noise to the exact values, the estimates shown in Table 1 are provided.

The resulting covariance matrix for the active subproblem is (same values result for the reactive one)

$$\operatorname{cov}\begin{bmatrix}\tilde{y}_b\\ \tilde{y}_i\end{bmatrix}_P = \operatorname{cov}\begin{bmatrix}P_{2e}\\ P_{3e}\\ P_{45}\\ P_{56}\end{bmatrix} = 10^{-5}\left[\begin{array}{cc|c} 15 & -10 & \\ -10 & 15 & -5 \\ \hline & -5 & 5 \\ & & 10 \end{array}\right].$$

As can be seen, the mutual terms coupling $\tilde{y}_b$ with $\tilde{y}_i$ are comparatively small or null in this case.

This example is useful to illustrate also a major advantage of the LSE over the conventional TSE, regarding the capability to individually handle all raw measurements. A standard SE, based on the bus-brunch model (B&B), relies on an auxiliary topology processor, aimed at identifying electrical buses and determining the net values of power measurements, as "seen" from the outer network. In this case, the topology processor could obtain the following active power "measurements" from actual raw measurements:

$$\begin{aligned} P_{2e} &= P_{12} - P_{23} \\ P_{3e} &= P_{23} - P_{34} \\ P_5 + P_6 &= -P_{16} - P_{45} \quad \text{(net power injection)} \end{aligned}$$

which means that six raw measurements are actually handled by the B&B model as three telemetered values (the same applies to the reactive power).

It should be noted that the local redundancy is relatively low in this example, which somewhat limits the possibility of properly identifying bad data at this level. For instance, if a 3.5% error is added to P_{12}^m the largest normalized residual is only 2.96 and the bad data flag is not triggered. On the other hand, when the added error is 4%, the largest normalized residual is 3.39 and bad data are detected. However, owing to the critical redundancy level of this example, both P_{12}^m and P_{61}^m share this abnormally high residual, which means that the bad data cannot be identified.

B. Transmission Substation: Nonlinear Integral Model

Extending the cluster notion to the entire substation (or even a few adjacent substations), and not only to the busbar sections of the same rated voltage, leads to a nonlinear model containing lossy elements. The advantage is that many more measurements can be locally processed, including phase angle differences provided by new generations of IEDs [55], increasing in this way the local redundancy. More detailed three-phase models could be even adopted if needed. The payoff is the added complexity of the nonlinear solution process, particularly when this constitutes the first stage of a two-stage TSE process.

The ideas will be illustrated with the help of the substation shown in Fig. 8, composed of two voltage levels ("linear substations" from the point of view of Section VII-A). For convenience, area 1 (right) will be again treated at the physical level, whereas the conventional B&B model will be adopted for area 2 (left). Fig. 10 represents the compact substation diagram, including one of the possible proper trees. The differences between the subtree corresponding to area 1 and the tree of Fig. 9 can be explained as follows (the reader can either skip these subtle details or see, for instance, [6, Ch. 8] for a deeper treatment).

- The external injections at buses 5 and 6 in the linear case become power transformers in this one. As nonzero impedance branches are links by definition, this implies that the CB branches 4–5 and 5–6 cannot be simultaneously links.
- A new branch at bus 4 (or at any other physical node within area 1) is added to take into account the null-injection constraints. These are needed to

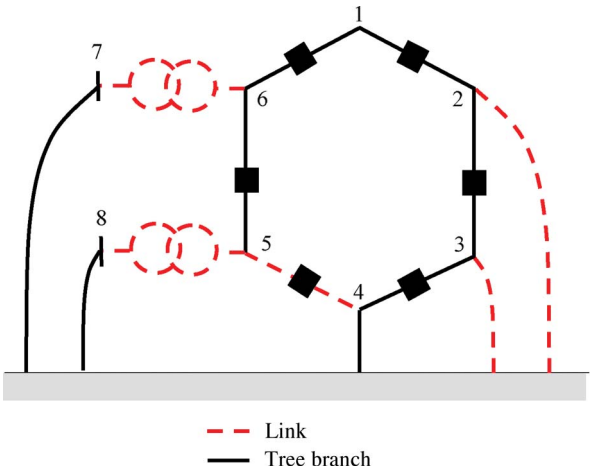


Fig. 10. *Graph associated with the substation of Fig. 8.*

assure that the physical subsystem composed of the CB ring is lossless (the sum of power flows through the four external branches is zero).

The state vector contains in this case a mixture of conventional and power flow variables (the latter required by the detailed physical model of area 1), as follows:

$$y = \begin{bmatrix} y_V \\ \hline y_P \\ y_Q \end{bmatrix} \quad \begin{cases} y_V^T = [V_1, V_7, V_8, \theta_7, \theta_8] \\ y_P^T = [P_{2e}, P_{3e}, P_{45}] \\ y_Q^T = [Q_{2e}, Q_{3e}, Q_{45}] \end{cases}$$

where the phase reference is taken at bus 1. This yields a total of 11 state variables, for which 26 measurements are available (six voltage magnitudes plus ten pairs of power flow measurements). In addition, nonlinear zero-injection constraints at bus 4 should be added.

The resulting measurement model comprises linear relationships, such as

$$P_{23}^m = P_{3e} + P_{45} + e_{p23}$$

as well as nonlinear ones

$$P_{76}^m = f_p(V_1, V_7, \theta_7) + e_{p76}$$

where $f_p(\cdot)$ is the well-known nonlinear function relating the active power flow through a transformer with the terminal bus voltages. The interested reader is encouraged to obtain the remaining expressions.

Notice that, unlike in the linear case, the variables in the subsets y_V, y_P, and y_Q become coupled in general by the presence of the nonlinear constraints.

From the perspective of the two-stage procedure, in which the substation is considered a component (Stage 1) of a larger TSO system (Stage 2), it is convenient to decompose the state vector as follows:

$$y = \begin{bmatrix} y_b \\ y_i \end{bmatrix}$$

where

$$y_b^T = [P_{2e}, P_{3e}, Q_{2e}, Q_{3e}, V_1]$$

and y_i contains the remaining variables (the internal ones).

For a measurement set which, regarding area 1, is identical to that of the linear substation case, the estimates corresponding to the state variables are shown in Table 2.

Table 2 Values of State Variables for the Nonlinear Substation Case

		Exact	Estimated
y_b	P_{2e}	1.3000	1.3296
	P_{3e}	1.1000	1.0755
	Q_{2e}	0.1300	0.1307
	Q_{3e}	0.1100	0.1232
	V_1	1.0200	1.0335
y_i	P_{45}	-1.2000	-1.1955
	Q_{45}	-0.1200	-0.1302
	V_7	1.0735	1.0851
	V_8	1.1099	1.1246
	θ_7	0.2322	0.2307
	θ_8	0.3169	0.3056

As stated earlier, the redundancy in the nonlinear case ($28/11 = 2.55$) is higher than in the linear one ($16/9 = 1.78$), which is an important aspect if topology errors or bad data have to be locally addressed. Referring again to the single bad datum analyzed in the linear substation case, affecting P_{12}^m, in this case even though the associated error is 3.5%, the largest normalized residual (3.45) exceeds the customary threshold 3, the next one in the ranking being 1.66. Therefore, such an error that could be barely detected in the linear case is safely identifiable when the entire substation is considered at once.

C. Distribution Substation and Associate Feeder System

Unlike transmission substations, in which very few lines and perhaps a lumped equivalent load are connected to each voltage level, distribution substations deliver power to a large amount of secondary transformers through a set of MV radial feeders (normally open switches allow backup service in case of a single failure).

As discussed in Section IV-D, the deployment of a new generation of sensors and meters, connected via relatively inexpensive communication channels with a DMS, in combination with well-established forecasting and data-mining techniques, is opening the way to the possibility of having a moderately redundant set of measurements and pseudomeasurements at the feeder level. In this upcoming context, the right tool capable of handling all these heterogeneous sources of information, according to their statistical quality, is the SE. In fact, SEs for distribution systems were considered long ago [57], when the required infrastructure was still far in the horizon.

In practice, up to 20 or more feeders may be connected to one or at most two MV busbar sections. As every feeder typically reaches dozens of secondary transformers (or even hundreds in rural dispersed areas), the size of the resulting SE problem may be discouraging. This section explains how the two-stage procedure can be used to take advantage of the weak electrical coupling among radial feeders.

A simplified distribution substation arrangement, with just two short feeders, is shown in Fig. 11 (the secondary

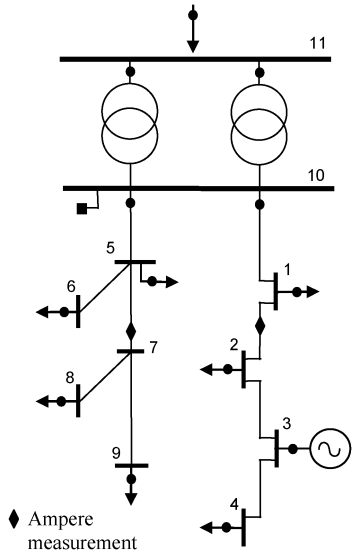


Fig. 11. *Sample distribution substation and associate feeders.*

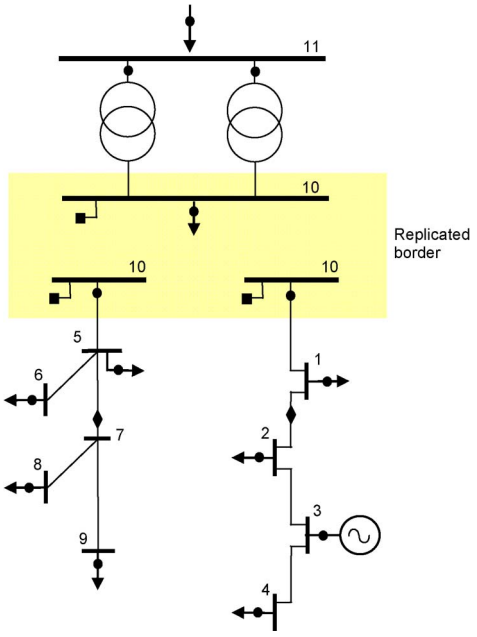


Fig. 12. *Resulting subsystems for the distribution substation.*

transformers at each feeder bus are not shown for simplicity). The feeder on the right, without laterals, is typically found in urban areas. It contains a DG, whose power injection exceeds in this example the load of the three remaining buses. While the power injected by the DG is accurately known, it is assumed that the load demand of buses 2 and 4, being provided by forecasting tools, is poorly defined. On the other hand, the feeder on the left corresponds with those of rural areas. In this case, a null-injection constraint (bus 7) coexists with power injections whose values are generally less accurate than those at transmission levels. Roughly in the middle of both feeders an Ampere measurement (provided for instance by a fault current detection device) is assumed. Typically, the measurement set within the substation is expected to be more accurate than the feeder measurements.

In this case, the border area allowing the overall SE problem to be decomposed into $f+1$ decoupled problems (f being the number of feeders) is simply the common bus from which the feeders hang, which is triplicated. Assuming the phase reference is taken for convenience at bus 10, the set of border variables reduces to V_{10} for all subproblems.

Fig. 12 shows the three decoupled systems that result in this example. The voltage magnitude measurement V_{10} can be used three times, once for each subsystem, but then its local weighting factor should be reduced accordingly (i.e., divided by three) so that its net effect, statistically speaking, is the same as if the global problem was solved at once. Similarly, the power flows at the head of both feeders can be used twice, once for each feeder and once more (in aggregated form, taking into account the null-injection constraint at bus 10) for the substation. Accordingly, their weighting coefficients should be divided by two.

As in any regular SE problem, the state vector for each subsystem comprises the voltage magnitudes and phase angles of all buses (the reference is in the border). Therefore, Stage 1 consists of solving three nonlinear systems of the form (17) by means of the iterative scheme (6). In turn, Stage 2 is composed of two linear systems, compactly represented by (18) and (19), which are coupled by the off-diagonal blocks of the gain matrices. In this simple case, the system (18) constraining the border variable reduces to

$$\begin{bmatrix} \tilde{V}_{10}^{(f1)} \\ \tilde{V}_{10}^{(f2)} \\ \tilde{V}_{10}^{(sub)} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} V_{10} + e_b.$$

Table 3 collects relevant results corresponding to the border magnitudes. The column under "Optimal" heading corresponds to the conventional solution of the entire problem, which is considered the optimal one. The minor differences that can be observed between "Optimal" and

Table 3 Values of Border Power Flows and Border Voltage for the Distribution Substation

	Exact	Optimal	Stage 2	$G_{ib}=0$
P_{10-1}	-0.1065	-0.1026	-0.1065	-0.1262
P_{10-5}	0.3819	0.3775	0.3763	0.3763
Q_{10-1}	-0.1204	-0.1213	-0.1266	-0.1398
Q_{10-5}	0.3910	0.3925	0.3897	0.3897
V_{10}	0.9913	0.9912	0.9912	0.9919

Stage 2" columns are due to the fact that Stage 2 has been run only once for this particular simulation. If the two-stage process was repeated, according to the complete methodology for the nonlinear–linear case (Section VI-D), the resulting differences between both columns would virtually vanish. Finally, the rightmost column shows the results that would be obtained by neglecting the off-diagonal blocks of G_1 (weighting matrix for Stage 2). As can be seen, the accuracy loss is not acceptable in this case, and presumably in all cases in which the coupling between border and interior variables is not negligible.

In this environment, both stages will run on the same computer, most likely located at the substation, which reduces the complexity associated with information exchanges between distant processors.

Upon completion of the two-stage local process, the border variables (in this case the power injections P_{11} and Q_{11} and the voltage magnitude V_{11}), along with the gain matrix arising during the last iteration, will be passed on to the TSO-level SE (the interior variables might also be involved if the block iterative scheme was adopted).

VIII. TSO-LEVEL STATE ESTIMATION

The remote data arriving from satellite LSEs (Stage 1), essentially composed of power flow and voltage magnitude estimates, as well as the associate statistical information, has to be integrated within the framework of the TSE (Stage 2). At this level, the resulting measurement model is mostly nonlinear, which involves the iterative solution of the system (7).

In this regard, the reader may recall that the design of the two schemes described in Section VI-D to solve Stage 2 in the linear case (reduced or iterative model) was based on the assumption that the number of clusters is moderate while the number of interior variables for each cluster is very high, which really calls for a distributed approach. However, the situation arising in the two-stage TSE is somewhat the opposite, namely the number of clusters j (i.e., substations) is very high, while the number of interior variables for each cluster is comparatively small (typically much lower than that of border variables, at least for transmission substations). Therefore, even though Stage 1 (substation level estimation) is performed in a distributed manner, the TSO-level coordination phase (Stage 2) should be better centralized, as the potential benefits of distributing the computations would be offset by the burden and complexity of exchanging information and coordinating the solution process with satellite processors.

Solving Stage 2 in a centralized manner provides the additional advantage that a standard SE can be resorted to, with minor adaptations. Major differences between the TSE that centrally performs Stage 2 and a conventional SE are [49] as follows.

- The two-stage TSE should be able to deal with nondiagonal yet sparse weighting matrices, characterizing estimates arriving from the substations. Indeed, conventional SEs customarily neglect coupling covariance terms between raw measurements provided by RTUs, which may not be acceptable in all cases [58]. Therefore, this should be a welcome addition. Anyway, as the weighting matrix is made up in this case of small diagonal blocks, one for each substation, this should not be a problem (in the linear substation case, each block is composed in turn of three decoupled blocks).
- A small set of additional state variables, corresponding to substation interior variables, must be accommodated into the state vector. Note however that this slightly augmented state vector would be anyway needed should the implicit GSE model be adopted [34].
- A reduced set of measurements, mostly composed of power flow and voltage magnitude measurements, arrives to Stage 2 (most injection measurements are locally processed). This may simplify to some extent certain auxiliary functions associated with the SE process.

It is worth noting that the computational effort of Stage 2 is virtually unaffected by the addition of more and more redundant measurements, as they are nearly always handled at the substation level (Stage 1). The relative computational efficiency of the two-stage SE increases therefore with the measurement redundancy. This will be particularly interesting in future "smart substations," with highly redundant measurement sets, where the proposed hierarchical approach will release the TSO-level SE from the burden of processing a huge number of raw measurements.

An added advantage of the two-stage scheme, arising from Stage 1 being implemented in a geographically distributed manner, is the reduction of the communication bandwidth requirements.

Other complementary functions, such as bad data processing, can also be partly performed in a distributed manner during Stage 1, provided the redundancy is high enough (this was illustrated in the previous section). Some bad data may have to be filtered out anyway at Stage 2.

In the following Section VIII-A and B, the TSE formulation will be separately addressed for the linear and nonlinear substation models considered above. A small example will be used to illustrate the ideas.

A. Linear Substation Model

As explained in Section VII-A, when a substation is decomposed into elemental subsystems, each associated with the physical equipment rated at the same voltage, the resulting local model is linear. The TSE model is then composed of as many linear clusters as electrical nodes, coupled in a nonlinear manner by external branches, including power transformers in this category. In this case, the two-stage linear–nonlinear methodology presented in

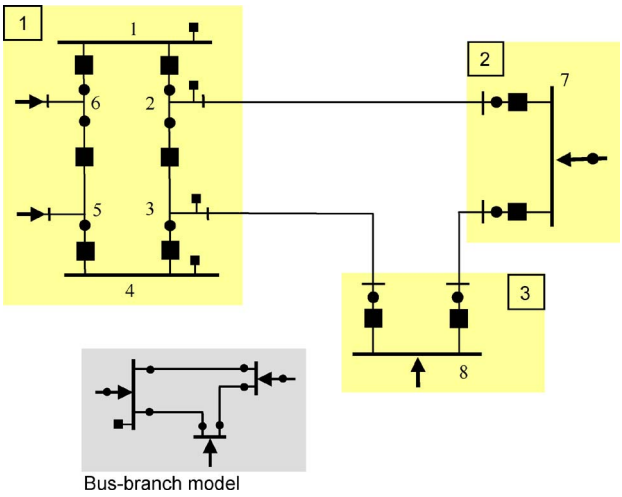


Fig. 13. *Sample transmission network with three linear substations.*

Section VI-B is directly of application. The role of the TSE is therefore to process in a centralized fashion, by solving (7), the prefiltered measurements provided by each cluster.

The sample system represented in Fig. 13 comprises three substations, one of them modeled in detail, as in Section VII-A. A conventional SE based on the B&B model would resort to the simplified diagram depicted at the bottom-left of the figure. In this case, the resulting measurements are generated during the topology processing stage by properly adding/subtracting raw measurements.

Table 4 shows, for a single snapshot, the simulation results corresponding to border power flows. From left to right, the table presents: 1) exact power flow values; 2) estimates provided by a generalized SE, capable of dealing with detailed substation models and, hence, with all raw measurements simultaneously (this is the optimal solution in the WLS sense); and 3) estimation errors corresponding to the conventional B&B model, Stage 1 and Stage 2, respectively, with respect to the optimal solution. In the last row, the mean of the absolute value errors (MAE) for the 12 power flow estimates is given. For the substation 1, the same measurements as in Section VII-A are used. In consequence, the results corresponding to Stage 1 are in agreement with those shown in Table 1 for y_b. Also, as Stage 1 is linear, the results provided by Stage 2 are optimal in this case after a single run (no need to recompute Stage 1 gain matrices). The results provided by the classical B&B model are suboptimal, but better than those after Stage 1. This is so because the B&B model, based on a subset of measurements generated by the topology processor, does not take full advantage of the redundancy that can be achieved when the detailed substation model is used. Moreover, as the raw measurements in substation 3 are locally critical, no possibility of local prefiltering exists in this case (Stage 1 is skipped for this substation).

B. Nonlinear Substation Model

In this case, each substation, including power transformers, is considered an indivisible entity, as in Section VII-B. The TSE model is then composed of as many nonlinear clusters as substations (several adjacent substations could be merged for convenience in a single cluster), coupled in a nonlinear manner by external branches (lines only). Note that transformerless substations reduce anyway to linear clusters. In this context, the fully nonlinear two-stage methodology presented in Section VI-A can be applied.

The system represented in Fig. 14 differs from that of Fig. 13 in substation 1, which now contains two power transformers (this affects also the B&B model, as shown at the bottom left).

Table 5 is the counterpart of Table 4. For substation 1, the same measurement data as in Section VII-B are used. Consequently, the results corresponding to Stage 1 are in agreement with those shown in Table 2 for y_b. In this case, it makes sense to repeat Stages 1 and 2 by updating the gain matrix of the only nonlinear cluster (substation 1). Interestingly, the results of this second run (estimation errors shown at the rightmost column) are hardly distinguishable

Table 4 Simulation Results for the Three-Bus System of Fig. 13

	Exact	Optimal	Estimation error (p.u.)		
			B & B	Stage 1	Stage 2
P_{12}	1.3000	1.2897	-4.9E-03	4.0E-02	4.0E-10
P_{13}	1.1000	1.0959	-1.0E-02	-2.6E-02	1.2E-10
Q_{12}	0.1300	0.1291	-5.8E-04	-2.6E-03	3.3E-11
Q_{13}	0.1100	0.1114	-8.1E-04	1.6E-03	4.2E-10
P_{21}	-1.2925	-1.2823	4.9E-03	1.7E-02	-3.6E-10
P_{23}	-0.5021	-0.4940	-3.3E-03	-1.1E-02	-3.2E-10
Q_{21}	-0.0923	-0.0918	2.9E-04	-1.1E-03	1.2E-10
Q_{23}	-0.0358	-0.0338	2.0E-05	2.3E-03	4.0E-10
P_{31}	-1.0958	-1.0917	9.9E-03		-1.0E-10
P_{32}	0.5031	0.4949	3.4E-03		3.3E-10
Q_{31}	-0.0889	-0.0903	4.3E-04		-3.4E-10
Q_{32}	0.0408	0.0386	7.0E-05		-3.8E-10
MAE			3.2E-03	1.3E-02	2.8E-10

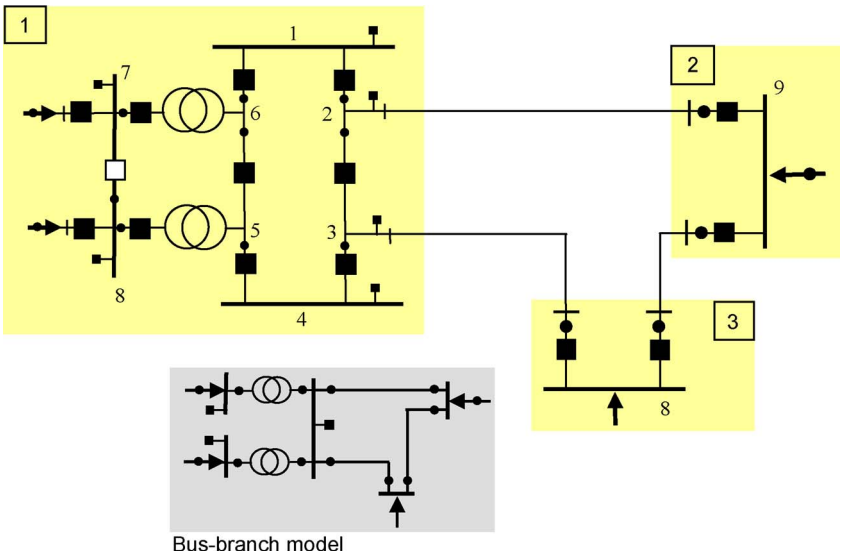


Fig. 14. ***Sample transmission network with a nonlinear substation.***

from those of the first one, and both can be deemed optimal for any practical purpose. Only when the measurement set contains unusually high noise levels is the second run justified. In the last row, the mean of the absolute value errors (MAE) for the 12 power flow estimates is given.

The results shown in Table 5 refer to a single snapshot. Repeating the experiment 1000 times provides a clearer picture of the resulting estimation accuracy for each SE model (random noise with $\sigma = 0.01$ is used to generate all measurements). Fig. 15 shows the probability density function (pdf) of the average estimation errors corresponding to the six pairs of border power flows. As expected, the conventional B&B model is more accurate than Stage 1 but less than Stage 2, according to the resulting redundancy levels (the pdf corresponding to the optimal solution is indistinguishable from that of Stage 2, first run).

In this context, it may be necessary to partly solve (7) in a distributed manner, for those more complex substations containing a large number of internal variables. This may apply, for instance, to distribution substations when feeders are involved in the SE process (Section VII-C).

IX. MULTI-TSO REGIONAL STATE ESTIMATION

The factorization-based approach described in Section VI-A provides also the theoretical basis for an upper SE hierarchy, in which the set of interconnected TSEs constitutes the first level and the RSE is the second level.

The geographical decomposition is achieved in a natural manner by replicating the state variables corresponding to tie-line terminal buses, which are shared by adjacent TSOs. In this context, both SE levels lead in general to nonlinear models, unless all raw measurements are processed at least once at the TSO level, which is generally the case. When this happens, the RSE solution reduces to a

Table 5 Simulation Results for the System of Fig. 14

	Exact	Optimal	Estimation error (p.u.)			
			B & B	Stage 1	Stage 2(I)	Stage 2(II)
P_{12}	1.3000	1.2937	-1.1E-02	3.6E-02	6.7E-07	-3.0E-08
P_{13}	1.1000	1.1025	-2.0E-02	-2.7E-02	9.4E-07	-4.2E-08
Q_{12}	0.1300	0.1317	-5.3E-03	-2.5E-03	-2.8E-06	-1.8E-07
Q_{13}	0.1100	0.1166	-8.3E-03	4.1E-03	-5.2E-06	-2.6E-07
P_{21}	-1.2925	-1.2864	1.1E-02	2.1E-02	-5.1E-07	2.8E-08
P_{23}	-0.5021	-0.4925	-5.3E-03	-1.3E-02	7.1E-08	-6.4E-09
Q_{21}	-0.0923	-0.0952	4.4E-03	3.8E-03	3.6E-06	1.7E-07
Q_{23}	-0.0358	-0.0326	-1.5E-03	5.6E-04	-1.2E-06	-3.1E-08
P_{31}	-1.0958	-1.0984	2.0E-02		-8.5E-07	4.0E-08
P_{32}	0.5031	0.4935	5.3E-03		-5.1E-08	6.2E-09
Q_{31}	-0.0889	-0.0958	7.5E-03		5.7E-06	2.5E-07
Q_{32}	0.0408	0.0372	1.6E-03		1.3E-06	3.0E-08
MAE			8.4E-03	1.3E-02	1.9E-06	8.9E-08

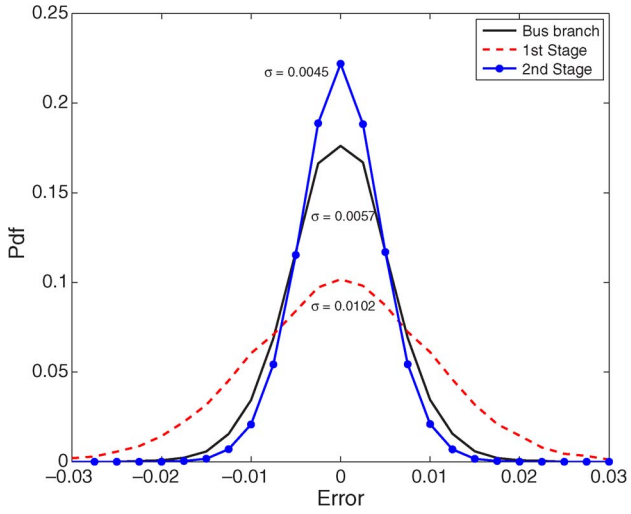


Fig. 15. ***Probability density function of average estimation errors.***

linear system in which the nonlinearity of the overall SE model is embedded into the weighting matrix, composed of the gain matrices provided by each TSE.

In the following, it will be assumed that the RSE model (Stage 2) is linear and that the blocked iterative solution scheme (25)–(26) described in Section VI-D is adopted. This increases the complexity of the interactions between each TSE and the RSE but has the advantage that all internal variables y_i are locally updated as a byproduct. Under these assumptions, main features of the RSE are as follows.

- Each TSO submits to the RSE its estimate of the border variables ($\tilde{y}_b$), along with the statistical information associated with the interior components, represented in compact form by the auxiliary vector (27). The overlapping degree determines the amount of border variables to be exchanged, but this is usually very small in relative terms. This is a welcome feature in nowadays deregulated systems, in which confidentiality of data is a major concern.
- The state vector at the RSE level comprises the border variables shared by the interconnected areas, as well as an auxiliary vector u composed of relative phase angle references, intended to synchronize all areas with respect to a global phase origin. Note that this was not an issue for the other applications of the two-stage procedure considered in this paper. In the distribution substation case, this is so because the common bus is chosen as the phase origin. At the TSO level phase angles are simply not included in the estimates provided by substations (in the presence of PMUs this may not be the case, as discussed in the next section).
- For each border component of the state vector two or more estimates are available in general, which increases the redundancy to estimate the auxiliary vector u.

A majority of multiarea SE procedures published up to date, combining heuristics and/or relaxation techniques with two-level WLS-based optimization, can be considered simplified particular cases of the general framework discussed herein, frequently providing suboptimal solutions [52], [53].

As an example, consider the system represented in Fig. 16, in which three IEEE 14-bus systems are interconnected by six tie-lines. The two scenarios reported in [50] will be summarized below.

A. Gaussian Errors

Two measurement sets with Gaussian errors are tested as follows.

- Set 1: containing the voltage magnitudes and the power injections at all the nodes, plus the power flows at both sides of all branches. This yields a total of 390 measurements and a redundancy of 4.64, for the entire 3-TSO system.
- Set 2: same as 1, except that branch power flows are measured at one terminal only. This yields a total of 258 measurements and a redundancy of 3.07.

The standard deviations associated with the measurements are: 0.01 for voltage magnitudes, 0.02 for power injections, and 0.015 for power flows. In the tests presented below, the voltage magnitude of border buses and tie-line power flow measurements are shared by neighboring TSOs, for which the respective weighting factors are suitably scaled.

Table 6 shows the 200-run average of the MAEs (taking exact values as references) associated with voltage magnitudes and phase angles for the complete set of state variables. The first row refers to the global SE solution, as if the entire system was a single TSO (optimal estimate). Two incomplete solutions of Stage 2 are added for comparison: $[1b-1i]$ stands for a single iteration of the block iterative scheme (25)–(26), in which both border and interior variables are updated only once after Stage 1; $[1b-0i]$ refers to the simplest possible arrangement, in which only border variables are updated.

As can be seen, even a single iteration of Stage 2 may provide sufficient estimation accuracy for practical purposes. Updating the interior variables at least once is however strongly recommended, particularly near the border. After four iterations, Stage 2 converges virtually to the reference solution, for a convergence threshold of 0.0001. As expected, more accuracy is obtained for measurement set 1.

B. Border Bad Data

The bad data detection and identification process is customarily implemented through the largest normalized residual ($|r_N|$) technique. For this purpose, several

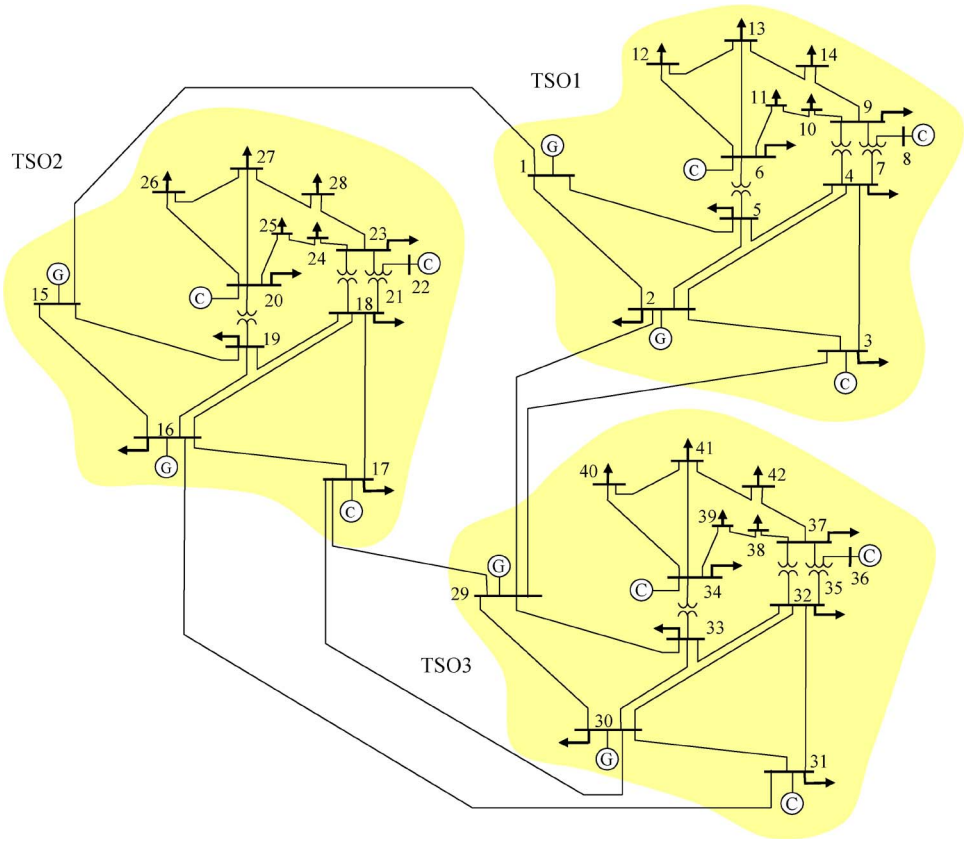


Fig. 16. ***Regional system composed of three 14-bus interconnected systems.***

scenarios are generated incorporating a single bad datum (10 σ) to measurements in the border area, the remaining ones (same two sets as before) being assumed as "exact." This way, the influence of random factors in the results is eliminated. Fig. 17 shows the five candidate bad data sequentially tested, all of them around the border bus 1.

Table 7 collects the two largest values of $|r_N|$ for each scenario, as well as the associate measurements, that result after the local TSO solution (Stage 1). Separate values are provided for the two involved partners (TSO 1 and TSO 2). It can be concluded that for the highly redundant

Table 6 Mean Absolute Value Errors for the Set of Border State Variables

Solution scheme	Meas. Set 1		Meas. Set 2	
	V	θ	V	θ
Optimal	0.00172	0.00136	0.00287	0.00227
[1*b*-1*i*]	0.00188	0.00137	0.00297	0.00229
[1*b*-0*i*]	0.00209	0.00138	0.00303	0.00228

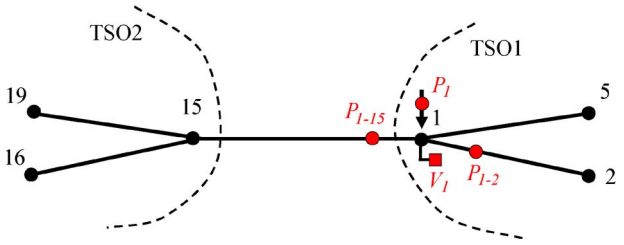


Fig. 17. ***Tested bad data around tie-line 1–15.***

Table 7 Two Largest $|r_N|$ for a Single Bad Datum After Stage 1

TSO 1				
Bad data	Meas. Set 1		Meas. Set 2	
	Max $\|r_N\|$	(meas.)	Max $\|r_N\|$	(meas.)
V_1	6.9213	V_1	6.8949	V_1
	0.4022	V_5	0.4991	$Q_{5,6}$
$P_{1,15}$	5.6626	$P_{1,15}$	4.4096	$P_{1,15}$
	3.1480	$P_{15,1}$	4.4059	P_1
P_1	7.1053	P_1	5.8763	P_1
	3.3575	$P_{1,15}$	5.8713	$P_{1,15}$
$P_{1,2}$	8.1889	$P_{1,2}$	6.5924	$P_{1,2}$
	3.8662	$P_{2,1}$	4.3346	P_1
TSO 2				
Bad data	Meas. Set 1		Meas. Set 2	
	Max $\|r_N\|$	(meas.)	Max $\|r_N\|$	(meas.)
V_1	6.9155	V_1	6.8732	V_1
	0.3717	V_19	0.7840	$Q_{1,15}$
$P_{1,15}$	5.6586	$P_{1,15}$	4.3957	$P_{1,15}$
	3.1599	$P_{15,1}$	4.3914	P_{15}

measurement set 1 the five bad data are correctly identified by both TSOs. However, with set 2 the bad data are detected but the identification is risky at least for $P_{1,15}$ and P_1.

Then, the first iteration of Stage 2 is partly performed, by updating only the border variables $(1b-0i)$, and the normalized residuals are computed again. In this case, all bad data are correctly identified by both TSOs.

Similar conclusions are obtained with not so extreme error levels. For instance, bad data on $P_{1,15}$ and P_1, when generated with 4σ, cannot be detected after Stage 1 and are barely identified after the first half iteration of Stage 2 $(1b-0i)$.

Therefore, bad data in tie-line and border bus measurements can be safely handled only at the RSE level (Stage 2), unless the redundancy is very high or the subset of measurements adjacent to the bad data is quite accurate. Indeed, this may be the case in the presence of PMUs, whose role in this multilevel paradigm will be discussed in Section X.

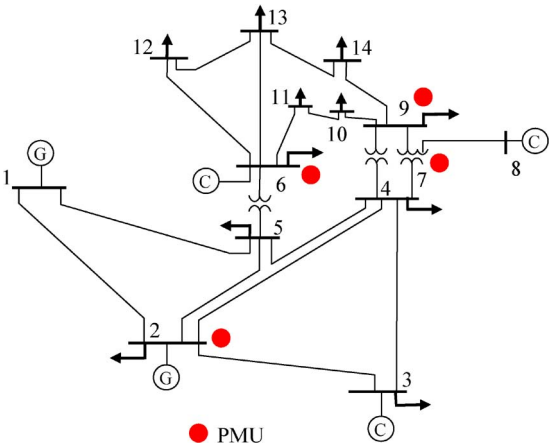


Fig. 18. *Optimal PMU placement for full network observability.*

X. INCORPORATION OF PMUs

Phasor measurements provide phase angles of bus voltage and branch current phasors that have long been missing from measurement sets since the introduction of state estimation function in the control centers [36]. Availability of direct measurement of these phase angles necessitates some changes in the state estimation formulation and also opens up new opportunities of enhancement for existing SEs. While PMUs are being deployed in increasing numbers all over the globe, their placement is commonly carried out in limited number of units at a time due to physical, operational, and financial constraints. Moreover, depending upon the number of channels they have, their placement strategy will have to be modified.

One simple yet fundamental difference in the formulation of the state estimation problem is the disappearance of the user assigned reference angle [59]. When collecting phasor measurements from PMUs at a PDC, one of the PMUs is typically chosen as a reference and all other measurements are reported with respect to this reference angle. This practice may however lead to erroneous results if the chosen reference is incorrect or missing. This problem can be easily overcome by using a reference-free SE where conventional as well as phasor measurements are processed without externally assigning any reference angle.

Given a limited budget and communication infrastructure, PMUs can be placed at strategic locations in order to enhance energy management functions. The strategy to place PMUs will depend both on the type of PMUs being considered as well as the overall placement objective and priorities. PMUs can be assigned to buses or branches based on their available channels [60]. Assuming no channel limits on PMUs, it can be shown that full network observability can be achieved by strategically placing such PMUs at about one third of the system buses [61]. Considering the small 14-bus system of Fig. 18, the four required PMUs are indicated by circles next to their optimal bus locations. It is further shown that increasing the channel numbers will rapidly saturate the benefits of having extra channels when the objective is full network observability [62].

As in the case of conventional measurements, PMU measurements can be used to improve bad data processing capability in SEs [63]. It can be shown that using relatively few PMUs all existing critical measurements can be transformed into redundant ones. As a result, all bad data in any measurement can be ensured to be detected making the measurement configuration statistically more robust.

An important benefit of having access to synchronized phasor measurements at certain locations is shown to be related to network parameter error identification [64]. Errors creep into network data bases over the years due to poor maintenance or user mishandling. Other causes include environmental effects such as temperature that changes certain parameters by statistically significant amounts. Such errors can be detected and identified by techniques that can be built into the state estimation solution. However, there may be limits to the capability to identify such errors when using only conventional measurements. Examples of such cases can be found in [64]. Placement of PMUs at strategic locations in order to avoid such deficiencies of the measurement design belongs to the set of unique benefits provided by PMUs for state estimation.

A. Local State Estimation

In the case of transmission substations, as indicated in Section VII-A and B, availability of phasor measurements will introduce benefits. In the linear case, an additional submodel can be considered using the phase angles of measured voltage phasors. Depending upon the location,

number, and type of phasor measurements available at the substation (both areas 1 and 2 of Fig. 8), the model may remain linear despite the lossy elements. This is similar to the situation where exclusive use of phasor measurements leading to a linear SE problem formulation for the conventional single stage case. Furthermore, having synchronized measurements (at least one from each area) will allow independent handling of the two areas of the substation that can then be synchronized based on the phasor measurements, before communicating the results up to the TSO level. Caution has to be exercised by not relying on a set of critical phasor measurements in case any of them carry bad datum and bias the results with no possible way of bad data detection. Given that most PMUs provide redundant voltage measurements at the substation, bad data detection will practically not be a major issue.

The same benefits translate to the case of distribution substation as long as similar level of PMU deployment exists. Unfortunately, current emphasis in deploying PMUs is on transmission systems and most distribution feeders of today probably are not equipped with PMUs.

B. TSO-Level State Estimation

Any available PMU measurements will augment the set of data arriving from satellite LSEs (Stage 1) to the TSO. Again, availability of synchronized measurements will provide two benefits at the TSO level. One is the added redundancy, which will help strengthen bad data processing capability at the TSO level and will also provide a direct way to merge LSE results even when boundary measurements are not sufficient or bad. The second advantage is dependent on the number, type, and location of phasor measurements, which may allow the nonlinear case of Section VIII-B to be transformed into a linear problem.

As implied by theory of network observability, conventional measurements have to be incident to boundary buses and/or tie-lines in order to enable merging of observable islands, which can be considered as locally processed substations or subsystems. PMU measurements, however, have the advantage that they can be placed at any point within the observable islands and still will facilitate mergers of their islands with other islands, irrespective of them being physical neighbors or not. Hence, Stage 2 of the multilevel SE paradigm will be greatly enhanced as transmission and distribution substations continue to be populated with more PMUs.

C. Multi-TSO/Regional Level State Estimation

At the regional level, the main concerns of coordination between TSOs have been articulated in several earlier publications [24], [65]–[67]. As in the case of TSO-level coordination among various substations, proper and efficient handling of binding constraints are crucial. These appear in form of measurements that are incident at boundary buses as well as various types of synchronized phasor measurements, which may be located anywhere in the operating areas supervised by individual TSOs. Having PMUs spread around the different TSOs allows flexible coordination schemes without explicitly relying on boundary measurements, which may not have the necessary local redundancy. In principle, having one phasor measurement per transmission system should suffice for synchronization of individual SE solutions. However, in practice, this may lead to poor performance due to bad or missing data. This issue can be addressed by ensuring that none of the PMU measurements are critical. There are also extreme conditions such as having a single current phasor measurement along the tie-line connecting two transmission systems, in which case multiple solutions will be possible for the combined solution. However, such cases are highly unlikely, given the multiplicity of channels available on typical PMUs.

XI. CONCLUSION

SEs have been around for some 40 years. In spite of the tremendous progress made in computational efficiency, numerical stability, modeling issues, robustness against bad data, etc., the structural design of the existing SEs essentially remains identical to that of the 1970s, based on a centralized EMS in charge of a single TSO network with minor or no interactions with neighboring or subordinate systems.

This paper proposes an alternative vision based on the anticipated new role of SEs in the context of the future smart grids. The paper first provides an overview of the recent advances in sensors as well as signal processing and communication technologies (PMUs, IEDs, communication protocols, etc.). It then goes on to propose a new paradigm based on a multilevel computation and communication architecture, which can sustain growth and complexity of data and information flow in various parts of future energy systems as they are monitored more closely and synchronously.

The proposed multilevel scheme is a generalization of existing bi-level schemes, both at the local and regional geographical levels. A theoretical framework, supporting the adequacy and optimality of the resulting hierarchical scheme, is provided. ■

APPENDIX

Starting from the optimality conditions of the conventional WLS SE, the two-stage factorized approach will be inferred for the most general, fully nonlinear case. This provides the mathematical justification of the particular implementations and main steps described in Section VI.

The first-order optimality conditions (FOOCs) of the conventional WLS SE are compactly given by [6]

$$H(\hat{x})^T W[z - h(\hat{x})] = 0. \tag{28}$$

In geometrical terms, the above expression implies that the estimated residual vector, when properly weighted with matrix W, must be orthogonal to the columns of the Jacobian matrix H. The optimal estimate $\hat{x}$ satisfying the above expression is obtained by iteratively solving the normal equation (2).

The factorized approach to solve the WLS problem arises when a vectory of intermediate variables is introduced so that

$$\left.\begin{array}{l} z = f_1(y) \\ y = f_2(x) \end{array}\right\} \quad \Rightarrow \quad h(x) = f_1[f_2(x)] \tag{29}$$

which, by the chain rule, leads to the following relationship between the Jacobians:

$$H(x) = F_1(y)F_2(x). \tag{30}$$

Taking into account (29) and (30), the FOOCs (28) can be written as

$$F_2^T F_1^T W\{z - f_1[f_2(\hat{x})]\} = 0 \tag{31}$$

where the dependence of the Jacobians on x has been omitted for simplicity of notation. Then, adding and subtracting the term $f_1(\tilde{y})$, where the meaning of $\tilde{y}$ will be made clear later, the following FOOCs are obtained:

$$F_2^T\left\{F_1^T W[z - f_1(\tilde{y})]\right\} + F_2^T F_1^T W\{f_1(\tilde{y}) - f_1[f_2(\hat{x})]\} = 0. \tag{32}$$

Moreover, assuming $\tilde{y}$ is sufficiently close to $f_2(\hat{x})$, the following linearization of $f_1(\cdot)$ can be performed:

$$f_1(\tilde{y}) - f_1[f_2(x)] \cong F_1[\tilde{y} - f_2(\hat{x})]. \tag{33}$$

Finally, replacing the above expression into (32) leads to

$$F_2^T\left\{F_1^T W[z - f_1(\tilde{y})]\right\} + F_2^T\left(F_1^T W F_1\right)[\tilde{y} - f_2(\hat{x})] = 0. \tag{34}$$

Therefore, the optimal estimate $\hat{x}$ in (28) can be alternatively obtained by finding the pair $\langle \tilde{y}, \hat{x} \rangle$ satisfying simultaneously the set of equations

$$F_1^T W[z - f_1(\tilde{y})] = 0 \tag{35}$$

$$F_2^T\left(F_1^T W F_1\right)[\tilde{y} - f_2(\hat{x})] = 0 \tag{36}$$

where it is important to realize that the Jacobian matrices F_1 and F_2 should be computed at the solution point $\langle \hat{y}, \hat{x} \rangle$, with $\hat{y} = f_2(\hat{x})$.

The solution to the original problem (28) can be then decomposed into two successive stages, each involving the solution of a WLS problem, which can be summarized as follows.

1) Stage 1: Obtain the estimate $\tilde{y}$ satisfying (35) by iteratively solving the associated normal equations

$$\left[f_1^T W F_1\right]\Delta y_k = F_1^T W[z - f_1(y_k)]. \tag{37}$$

2) Stage 2: Using the value $\tilde{y}$ provided by Stage 1 as "measurement" vector in (36), obtain the estimate $\hat{x}$ by iteratively solving the resulting normal equations

$$\left[f_2^T G_1 F_2\right]\Delta x_k = F_2^T G_1[\tilde{y} - f_2(x_k)] \tag{38}$$

where the "weighting" matrix $G_1 = F_1^T W F_1$ is the gain matrix of Stage 1.

During the first solution of Stage 1 the Jacobian F_1 is updated at each iteration, according to the current value y_k. However, when a new value $\hat{x}$ is provided by Stage 2, the Jacobian F_1 gets obsolete. Therefore, after the first run of Stages 1 and 2, the Jacobian F_1 should be recomputed with $\hat{y} = f_2(\hat{x})$ and Stage 1 run again, but this time keeping the Jacobian and gain matrices constant. Then, Stage 2 should be repeated with updated values of $\tilde{y}$, and so on until Stage 2 provides close enough values of $\hat{x}$ in two consecutive runs. In summary, in order to reach the optimal solution, there should be a sequence of Stage 1 and Stage 2 runs until full convergence, with constant Jacobian in Stage 1 as required by new $\hat{x}$ values provided by Stage 2.

An interesting particular case arises when Stage 1 is linear, the factorized model being

$$\left.\begin{array}{l} z = Ay \\ y = f_2(x) \end{array}\right\} \quad \Rightarrow \quad h(x) = Af_2(x). \tag{39}$$

In this case, the Jacobian F_1 becomes the constant matrix A, and a single execution of Stages 1 and 2 provides the optimal solution. Even in the fully nonlinear case, a single run of Stages 1 and 2 provides acceptable results for practical purposes, provided the raw measurements are sufficiently accurate.

REFERENCES

[1] R. N. Anderson, The Pew Center on Global Climate Change and the National Commission on Energy Policy, "The distributed storage-generation smart electric grid of the future," in *Proc. Workshop, The 10–50 Solution: Technol. Policies for a Low-Carbon Future*, 2004.

[2] B. Stott, O. Alsac, and A. J. Monticelli, "Security analysis and optimization," *Proc. IEEE*, vol. 75, no. 12, pp. 1623–1644, Dec. 1987.

[3] F. C. Schweppe, J. Wildes, and D. B. Rom, "Power system static state estimation—Part I, II, III," *IEEE Trans. Power App. Syst.*, vol. PAS-89, no. 1, pp. 120–135, Jan. 1970.

[4] F. C. Schweppe and E. J. Handschin, "Static state estimation in electric power systems," *Proc. IEEE*, vol. 62, no. 7, pp. 972–982, Jul. 1974.

[5] A. Monticelli, *State Estimation in Electric Power System. A Generalized Approach*. Norwell, MA: Kluwer, 1999.

[6] A. Abur and A. Gómez-Expósito, *Power System State Estimation: Theory and Implementation*. New York: Marcel Dekker, 2004.

[7] W. Tinney and R. Walker, "Direct solutions of sparse network equations by optimally ordered triangular factorization," *Proc. IEEE*, vol. 55, no. 11, pp. 1801–1809, Nov. 1967.

[8] A. Monticelli and A. Garcia, "Fast decoupled state estimators," *IEEE Trans. Power Syst.*, vol. 5, no. 2, pp. 556–564, May 1990.

[9] B. Stott, "Fast decoupled load flow," *IEEE Trans. Power App. Syst.*, vol. PAS-93, no. 3, pp. 859–869, May 1974.

[10] F. F. Wu, "Power system state estimation: A survey," *Int. J. Electr. Power Energy Syst.*, vol. 12, no. 2, pp. 80–87, Apr. 1990.

[11] M. Vempati, I. Slutsker, and W. Tinney, "Enhancements to givens rotations for power system state estimation," *IEEE Trans. Power Syst.*, vol. 6, no. 2, pp. 842–849, May 1991.

[12] A. Gjelsvik, "The significance of the lagrange multipliers in WLS state estimation with equality constraints," in *Proc. 11th Power Syst. Comput. Conf.*, Avignon, Aug. 1993, pp. 619–625.

[13] A. Gjelsvik, S. Aam, and L. Holten, "Hachtel's augmented matrix method—A rapid method improving numerical stability in power system static state estimation," *IEEE Trans. Power App. Syst.*, vol. PAS-104, no. 11, pp. 2987–2993, Nov. 1985.

[14] R. Nucera and M. Gilles, "A blocked sparse matrix formulation for the solution of equality-constrained state estimation," *IEEE Trans. Power Syst.*, vol. 6, no. 1, pp. 214–224, Feb. 1991.

[15] F. Alvarado and W. Tinney, "State estimation using augmented blocked matrices," *IEEE Trans. Power Syst.*, vol. 5, no. 3, pp. 911–921, Aug. 1990.

[16] A. Monticelli and F. F. Wu, "Network observability: Theory," *IEEE Trans. Power App. Syst.*, vol. PAS-104, no. 5, pp. 1042–1048, May 1985.

[17] G. R. Krumpholz, K. A. Clements, and P. W. Davis, "Power system observability: A practical algorithm using network topology," *IEEE Trans. Power App. Syst.*, vol. PAS-99, no. 4, pp. 1534–1542, Jul. 1980.

[18] K. A. Clements, "Observability methods and optimal meter placement," *Int. J. Electr. Power Energy Syst.*, vol. 12, no. 2, pp. 88–93, Apr. 1990.

[19] F. H. Magnago and A. Abur, "Unified approach to robust meter placement against bad data and branch outages," *IEEE Trans. Power Syst.*, vol. 15, no. 3, pp. 945–949, Aug. 2000.

[20] A. Monticelli and A. Garcia, "Reliable bad data processing for real-time state estimation," *IEEE Trans. Power App. Syst.*, vol. PAS-102, no. 5, pp. 1126–1139, May 1983.

[21] L. Mili, T. Van Cutsem, and M. Ribbens-Pavella, "Hypothesis testing identification: A new method for bad data analysis in power system state estimation," *IEEE Trans. Power App. Syst.*, vol. PAS-103, no. 11, pp. 3239–3252, Nov. 1984.

[22] M. K. Celik and A. Abur, "A robust WLAV state estimator using transformations," *IEEE Trans. Power Syst.*, vol. 7, no. 1, pp. 106–113, Feb. 1992.

[23] L. Mili, M. G. Cheniae, N. S. Vichare, and P. J. Rousseeuw, "Robust state estimation based on projection statistics of power systems," *IEEE Trans. Power Syst.*, vol. 11, no. 2, pp. 1118–1127, May 1996.

[24] T. Van Cutsem, J. L. Horward, and M. Ribbens-Pavella, "A two-level static state estimator for electric power systems," *IEEE Trans. Power App. Syst.*, vol. PAS-100, no. 8, pp. 3722–3732, Aug. 1981.

[25] J. M. Ruiz Muñoz and A. Gómez-Expósito, "A line-current measurement based state estimator," *IEEE Trans. Power Syst.*, vol. 7, no. 2, pp. 513–519, May 1992.

[26] A. Abur and A. Gómez-Expósito, "Detecting multiple solutions in state estimation in the presence of current magnitude measurements," *IEEE Trans. Power Syst.*, vol. 12, no. 1, pp. 370–375, Feb. 1997.

[27] A. Gómez-Expósito and A. Abur, "Generalized observability analysis and measurement classification," *IEEE Trans. Power Syst.*, vol. 13, no. 3, pp. 1090–1096, Aug. 1998.

[28] K. A. Clements, P. W. Davis, and K. D. Frey, "Treatment of inequality constraints in power system state estimation," *IEEE Trans. Power Syst.*, vol. 10, no. 2, pp. 567–574, May 1995.

[29] H. Singh, F. L. Alvarado, and W. Liu, "Constrained LAV state estimation using penalty functions," *IEEE Trans. Power Syst.*, vol. 12, no. 1, pp. 383–388, Feb. 1997.

[30] I. W. Slutsker, S. Mokhtari, and K. A. Clements, "Real time recursive parameter estimation in energy management systems," *IEEE Trans. Power Syst.*, vol. 11, no. 3, pp. 1393–1399, Aug. 1996.

[31] A. Monticelli, "Modeling circuit breakers in weighted least squares state estimation," *IEEE Trans. Power Syst.*, vol. 8, no. 3, pp. 1143–1149, Aug. 1993.

[32] O. Alsac, N. Vempati, B. Stott, and A. Monticelli, "Generalized state estimation," *IEEE Trans. Power Syst.*, vol. 13, no. 3, pp. 1069–1075, Aug. 1998.

[33] A. Monticelli, "Electric power system state estimation," *Proc. IEEE*, vol. 88, no. 2, pp. 262–282, Feb. 2000.

[34] A. de la Villa Jaén and A. Gómez-Expósito, "Implicitly constrained substation model for state estimation," *IEEE Trans. Power Syst.*, vol. 17, no. 3, pp. 850–856, Aug. 2002.

[35] Z. Jun and A. Abur, "Identification of network parameter errors," *IEEE Trans. Power Syst.*, vol. 21, no. 2, pp. 586–592, May 2006.

[36] A. G. Phadke, "Synchronized phasor measurements in power systems," *IEEE Comput. Appl. Power*, vol. 6, no. 2, pp. 10–15, Apr. 1993.

[37] A. G. Phadke and J. S. Thorp, *Synchronized Phasor Measurements and Their Applications*. New York: Springer-Verlag, 2008.

[38] A. Monticelli and F. Wu, "A method that combines internal state estimation and external network modeling," *IEEE Trans. Power App. Syst.*, vol. PAS-104, no. 1, pp. 91–103, Jan. 1985.

[39] J. D. McDonald, *Electric Power Substation Engineering*. Boca Raton, FL: CRC Press, 2003.

[40] [Online]. Available: www.energie-schweiz.ch/imperia/md/content/energiemrkteetrgertechniken/elektrizitt/strompanne03/12.pdf

[41] [Online]. Available: www.energy-regulators.eu/portal/page/portal/eer_home/eer_publications/ceer_ergeg_papers/electricity/2007/e06-bag-01-06_blackout-finalreport_2007-02-06.pdf

[42] *IEEE Standard Definition, Specification and Analysis of Systems Used for Supervisory Control, Data Acquisition, and Automatic Control*, IEEE Std. C37.1.1994, 1994.

[43] K. Brand, V. Lohmann, and W. Wimmer, *Substation Automation Handbook*. Bremgarten, Switzerland: Utility Automation Consulting Lohmann, 2003.

[44] *IEEE Standard for Synchrophasors for Power Systems*, IEEE Std C37.118-2005.

[45] I. H. Valenzo, "Information architecture design for the electricity distribution network," M.S. thesis, Economics of Infrastructures Section of the Technology, Policy and Management, Delft Univ. Technol., Delft, The Netherlands, 2009.

[46] [Online]. Available: http://www.energy-regulators.eu/portal/page/portal/EER_HOME/EER_INITIATIVES

[47] [Online]. Available: http://www.desertec.org/

[48] A. Bose, "Smart transmission grid applications and their supporting infrastructure," *IEEE Trans. Smart Grid*, vol. 1, no. 1, pp. 11–19, Jun. 2010.

[49] A. Gómez-Expósito and A. de la Villa Jaén, "Two-level state estimation with local measurement pre-processing," *IEEE Trans. Power Syst.*, vol. 24, no. 2, pp. 676–684, May 2009.

[50] European Community's 7th Framework Programme, "Algorithms for State Estimation of ETN," Deliverable 2.1 part 1, PEGASE Project (Pan European Grid Advanced Simulation and state Estimation), 2010.

[51] C. Gómez-Quiles, A. de la Villa Jaén, and A. Gómez-Expósito, "A factorized approach to WLS state estimation," *IEEE Trans. Power Syst.*, DOI: 10.1109/TPWRS.2010.2096830.

[52] T. Van Cutsem and M. Ribbens-Pavella, "Critical survey of hierarchical methods for state estimation of electric power systems," *IEEE Trans. Power App. Syst.*, vol. PAS-102, no. 10, pp. 3415–3424, Oct. 1983.

[53] A. Gómez-Expósito, A. de la Villa Jaén, C. Gómez-Quiles, P. Rousseaux, and T. Van Cutsem, "A taxonomy of multi-area state estimation methods," *Electric Power Syst. Res.*, DOI: 10.1016/j.epsr.2010.11.012.

[54] H. M. Kim, J. J. Lee, and D. J. Kang, "A platform for smart substations," in *Proc. Future Generat. Commun. Netw.*, 2007, vol. 2, pp. 579–582.

[55] A. de la Villa Jaén, P. Cruz Romero, and A. Gómez-Expósito, "Substation data validation by local three-phase generalized state estimators," *IEEE Trans. Power Syst.*, vol. 20, no. 1, pp. 264–271, Feb. 2005.

[56] A. Gómez Expósito and A. de la Villa Jaén, "Reduced substation models for generalized state estimation," *IEEE Trans. Power Syst.*, vol. 16, no. 4, pp. 839–846, Nov. 2001.

[57] C. N. Lu, J. H. Teng, and W.-H. E. Liu, "Distribution state estimation," *IEEE Trans. Power Syst.*, vol. 10, no. 1, pp. 229–240, Feb. 1995.

[58] E. Caro, A. Conejo, and R. Mínguez , "Power system state estimation considering measurement dependencies," *IEEE Trans. Power Syst.*, vol. 24, no. 4, pp. 1875–1885, Nov. 2009.

[59] J. Zhu and A. Abur, "Effect of phasor measurements on the choice of reference bus for state estimation," in *Proc. IEEE PES General Meeting*, Tampa, FL, Jun. 24–28, 2004, DOI: 10.1109/PES.2007.386175.

[60] B. Xu and A. Abur, "Observability analysis and measurement placement for system with PMUs," in *Proc. IEEE PES Power Syst. Conf. Expo.*, New York, Oct. 10–13, 2004, vol. 2, pp. 943–946.

[61] B. Xu, J. Y. Yeo, and A. Abur, "Optimal placement and utilization of phasor measurements for state estimation," in *Proc. Power Syst. Comput. Conf.*, Liege, Belgium, Aug. 22–26, 2005, pp. 1550–1555.

[62] M. Korkali and A. Abur, "Placement of PMUs with channel limits," in *Proc. IEEE Power Energy Soc. General Meeting*, Calgary, CA, Jul. 26–30, 2009, DOI: 10.1109/PES.2009.5275529.

[63] J. Chen and A. Abur, "Placement of PMUs to enable bad data detection in state estimation," *IEEE Trans. Power Syst.*, vol. 21, no. 4, pp. 1608–1915, Nov. 2006.

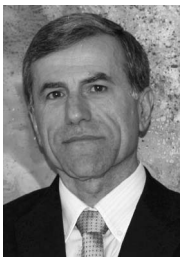
[64] J. Zhu and A. Abur, "Identification of network parameter errors using phasor measurements," in *Proc. IEEE Power Energy Soc. General Meeting*, Calgary, CA, Jul. 26–30, 2009, DOI: 10.1109/PES.2009.5275490.

[65] W. Jiang, V. Vittal, and G. T. Heydt, "Diakoptic state estimation using phasor measurement units," *IEEE Trans. Power Syst.*, vol. 23, no. 4, pp. 1580–1589, Nov. 2008.

[66] L. Zhao and A. Abur, "Multiarea state estimation using synchronized phasor measurements," *IEEE Trans. Power Syst.*, vol. 20, no. 2, pp. 611–617, May 2005.

[67] J. Weiqing, V. Vittal, and G. T. Heydt, "A distributed state estimator utilizing synchronized phasor measurements," *IEEE Trans. Power Syst.*, vol. 22, no. 2, pp. 563–571, May 2007.

ABOUT THE AUTHORS



Antonio Gómez-Expósito (Fellow, IEEE) received the Industrial Engineering degree in electrical engineering and the Doctor Engineering degree, both with honors, from the University of Seville, Seville, Spain, in 1982 and 1985, respectively.

Currently, he is a Full Professor at the University of Seville, where he is chairing the Department of Electrical Engineering, the Endesa Red Professorship, and the Electrical Energy Systems Post-Graduate Program. In addition to nearly 250 technical publications, he has coauthored several textbooks and monographs about circuit theory and power system analysis, among which *Power System State Estimation: Theory and Implementation* (New York: Marcel Dekker, 2004) and *Electric Energy Systems: Analysis and Operation* (Boca Raton, FL: CRC Press, 2008), stand out.

Dr. Gómez-Expósito is a Fellow of the IEEE for his work on power systems analysis and operation, and an IEEE/PES Distinguished Lecturer. He serves on the Editorial Board of the IEEE TRANSACTIONS ON POWER SYSTEMS starting in 2011.

Ali Abur (Fellow, IEEE) received the B.S. degree from Orta Doğu Teknik Üniversitesi, Ankara, Turkey, in 1979 and the M.S. and Ph.D. degrees from the Ohio State University, Columbus, in 1981 and 1985, respectively.

He was a Professor at Texas A&M University until November 2005 when he joined the faculty of Northeastern University, Boston, MA, as a Professor and Chair of the Electrical and Computer Engineering Department. He coauthored the book *Power System State Estimation: Theory and Implementation* (New York: Marcel Dekker, 2004), contributed to several book chapters, and published widely in IEEE journals and conferences. His research and educational activities have been in the area of power systems.

Dr. Abur is a Fellow of the IEEE for his work on power system state estimation. He serves on the Editorial Board of the IEEE TRANSACTIONS ON POWER SYSTEMS and POWER ENGINEERING LETTERS. He is an IEEE/PES Distinguished Lecturer.

Antonio de la Villa Jaén was born in Riotinto, Spain, in 1960. He received the electrical and doctor engineering degrees from the University of Seville, Seville, Spain, in 1983 and 2001, respectively.

Currently, he is an Associate Professor at the Department of Electrical Engineering, University of Seville. His primary areas of interest are computer methods for power system state estimation problems.

Catalina Gómez-Quiles (Student Member, IEEE) received the Eng. degree from the University of Seville, Seville, Spain, in 2006 and the M.Eng. degree from McGill University, Montreal, QC, Canada, in 2008, both in electrical engineering. She is currently working towards the Ph.D. degree in power system state estimation at University of Seville.

Her research interests include risk assessment in competitive electricity markets and mathematical and computer models for power system state estimation.